

## THESIS / THÈSE

### MASTER EN SCIENCES INFORMATIQUES

#### Implémentation d'une version interactive pour l'analyse de régression multiple

Hallo, Maria

*Award date:*  
1984

*Awarding institution:*  
Universite de Namur

[Link to publication](#)

#### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

#### Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

FACULTES UNIVERSITAIRES NOTRE - DAME DE LA PAIX - NAMUR

INSTITUT D'INFORMATIQUE

IMPLEMENTATION D'UNE VERSION INTERACTIVE  
POUR L'ANALYSE DE REGRESSION MULTIPLE

Mémoire présenté par

Maria Hallo

en vue de l'obtention du grade de  
Licencié et Maître en Informatique

ANNEE ACADEMIQUE 1983-1984

J'exprime ma gratitude envers Mme. M. Noihomme qui a accepté la direction de ce mémoire et que par ses nombreuses remarques a permis de mener à bien ce travail.

Je tiens aussi à remercier Messieurs P. Dagnelie et J. Claustriau du Centre de Recherches Agronomiques de l'Etat à Gembloux qui m'ont permis de recueillir le matériel bibliographique nécessaire pour l'élaboration de ce travail.

Mes remerciements vont aussi à Monsieur R. Palm pour ses explications concernant l'analyse de la régression multiple.

Je suis également reconnaissant envers Monsieur J. Barreto pour ses critiques et conseils.

En fin, Je remercie Mme. J. Bernard pour la mise en page et la dactylographie du texte, ainsi que tous ceux qui ont collaboré pendant l'élaboration de ce travail et en général durant mon séjour en Belgique.



## TABLE DE MATIERES

### CHAPITRE I. INTRODUCTION

### CHAPITRE II. ETUDE D'OPPORTUNITE.

pg.

2.1	Analyse de l'existant.	33
2.1.1	Diagramme de flux du système	3
2.1.2	Description des programmes	5
2.2	Estructure actuelle des données	14
2.2.1	Dictionnaire des données	14
2.2.2	Estructure d'accès	16
2.3	Critique de l'existant	19
2.4	Expression des besoins	19
2.5	Ebauche de solutions	20
2.6	Choix d'une estructure des données	20

### CHAPITRE III. ANALYSE DE LA REGRESSION MULTIPLE

3.1	Introduction	25
3.2	Terminologie et definitions	26
3.3	Modèle et hypothèse	31
3.3.1	Les conditions d'application	32
3.4	Estimation des paramètres	33
3.4.1	Téchniques de calcul	36
3.5	Choix des variables explicatives	39
3.5.1	Comparaison de toutes les équations de régression possibles.	40
3.5.2	Sélection régressive	40
3.5.3	Sélection progressive	41
3.5.4	Méthode de sélection mixte ou de régression pas à pas	42
3.6	Test de signification de coefficients de régression partielle	43
3.7	Description générale de solutions	45
3.7.1	Détails de calcul	46
3.8	Description de l'algorithme pour la méthode de régression pas à pas	48

### CHAPITRE IV. IMPLEMENTATION

4.1	Introduction	50
4.2	Module de gestion de menus	54
4.3	Modules de traitements de données	
4.3.1	Module de chargement des variables	56
4.3.2	Module de suppression des variables	59
4.3.3	Module de transformation des variables	61
4.3.4	Module d'impression des variables	62
4.3.5	Module d'accès aux structures de données	64
4.4	Module de régression multiple	65



## TABLE DE MATIERES

### CHAPITRE V. MANUEL DE L'UTILISATEUR

5.1	Introduction	68
5.2	Exécution du programme RMUL	69
5.3	Information sur le programme RMUL	70
5.4	Menu principale	71
5.5	Chargement des variables	72
5.6	Suppression des variables	73
5.7	Impression des variables	74
5.8	Transformation des variables	75
5.9	Régression multiple	76
5.10	Exemple d'utilisation	77
5.11	Analyse de résultats	101
		102

### CONCLUSIONS

### BIBLIOGRAPHIE

104

# IMPLEMENTATION D'UNE VERSION INTERACTIVE POUR L'ANALYSE DE REGRESSION MULTIPLE

## Chapitre 1 : INTRODUCTION

La régression multiple a pour objectif l'étude de la relation qui peut exister entre une variable observée appelée "dépendante" ou à expliquer et deux ou plusieurs variables explicatives ou "prédicteurs".

Le but de ce mémoire est d'étudier les aspects liés à l'implémentation d'une version interactive pour effectuer la régression multiple, destinée à être utilisée dans un centre de recherche agronomique. Actuellement cette application fait partie d'une bibliothèque de programmes statistiques qui est utilisée en mode batch.

Dans l'histoire de l'informatisation des techniques statistiques la plupart des travaux appartiennent à une des trois catégories suivantes :

- Elaboration de programmes simples
- Rassemblement d'algorithmes
- Systèmes statistiques.

Grand nombre de statisticiens arrivent à l'informatique en écrivant des programmes simples pour faire une analyse déterminée; de cette façon, chaque programme peut être écrit séparément pour chaque application avec l'inconvénient que plusieurs programmes ne peuvent pas être combinés séparément pour construire des analyses plus générales.

Une nouvelle approche apparaît avec l'idée de créer des sous-programmes qui réalisent des tâches spécifiques. On a développé une large collection de tels sous-programmes bien désignés et testés en laissant la tâche aux statisticiens de les utiliser dans leurs programmes. Dès lors, on simplifie la tâche de programmation et on améliore la qualité des programmes résultants.



Dans le cas des systèmes statistiques on a des collections de programmes qui essaient de donner à l'utilisateur une variété d'analyses statistiques d'une façon coordonnée et consistante; dans la plupart des cas, ils possèdent un langage spécial grâce auquel l'utilisateur spécifie l'analyse à réaliser.

Que ce soit par septicisme ou méconnaissance de l'informatique, les développements du software n'ont été acceptés que lentement par les concepteurs de programmes statistiques. Quelques systèmes ont été lents à utiliser des techniques statistiques nouvelles ou à ajouter des facilités informatiques, tels que la programmation interactive ou l'affichage graphique (1).

En considérant les aspects cités ci-avant, la première partie de ce travail consiste à faire une étude globale d'un système statistique utilisé en mode batch, ce qui se fait au chapitre 2.

Le chapitre 3 présente l'étude de la régression multiple : les conditions d'application, les méthodes de choix des variables et d'estimation des paramètres qui interviennent dans l'équation de régression pour un modèle linéaire. La méthode de régression pas à pas est décrite en détail.

Le chapitre 4 présente la description du logiciel pour faire l'analyse de la régression multiple et les traitements des données qui peuvent être utilisées dans le futur pour implémenter d'autres analyses statistiques.

Le chapitre 5 contient le manuel de l'utilisateur.

Et enfin, le chapitre 6 contient un exemple d'utilisation du programme et l'analyse des résultats.

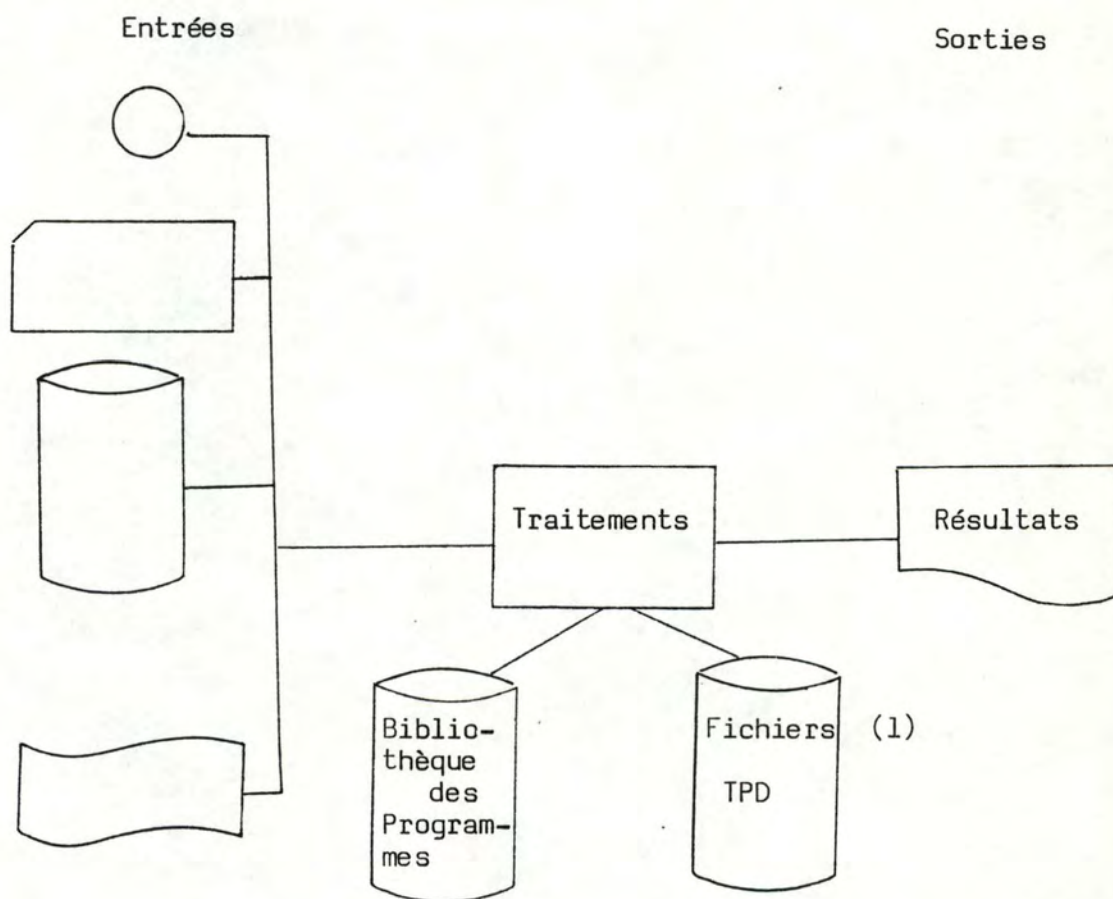


## Chapitre 2 : ETUDE D'OPPORTUNITE

Dans ce chapitre on analyse les aspects liés à la transformation en mode interactif d'une bibliothèque des programmes statistiques fonctionnant en mode batch.

### 2. 1. Analyse de l'existant

#### 2. 1. 1. Diagramme de flux du système



(1) Fichiers des titres, paramètres et données des variables  
(cf 2.2.3.)

### Entrées :

Dépendent du type de traitement à effectuer, on peut avoir :

- Spécification du type de traitement.
- Noms des fichiers où se trouvent les données.
- Spécifications des données concernées dans le traitement.
- Spécification des paramètres additionnels.

### Sorties :

On a comme sortie le résultat des traitements qui peuvent être stockés sur disque ou être imprimés sur papier.

### Données permanentes :

- Fichiers contenant les programmes que réalisent les traitements.
- Et, pour chaque utilisateur, les fichiers des titres, les paramètres et données des variables. Les paramètres donnent des informations sur les données.

### Traitements :

Comprend une série des programmes utilisés pour l'analyse statistique, parmi lesquels les plus utilisés dans la recherche agronomique ont été groupés en 4 types :

- Programmes généraux.
- Programmes de corrélation et de régression.
- Programmes d'analyse de la variance.
- Programmes d'ajustement.

## 2. 1. 2. Description des programmes

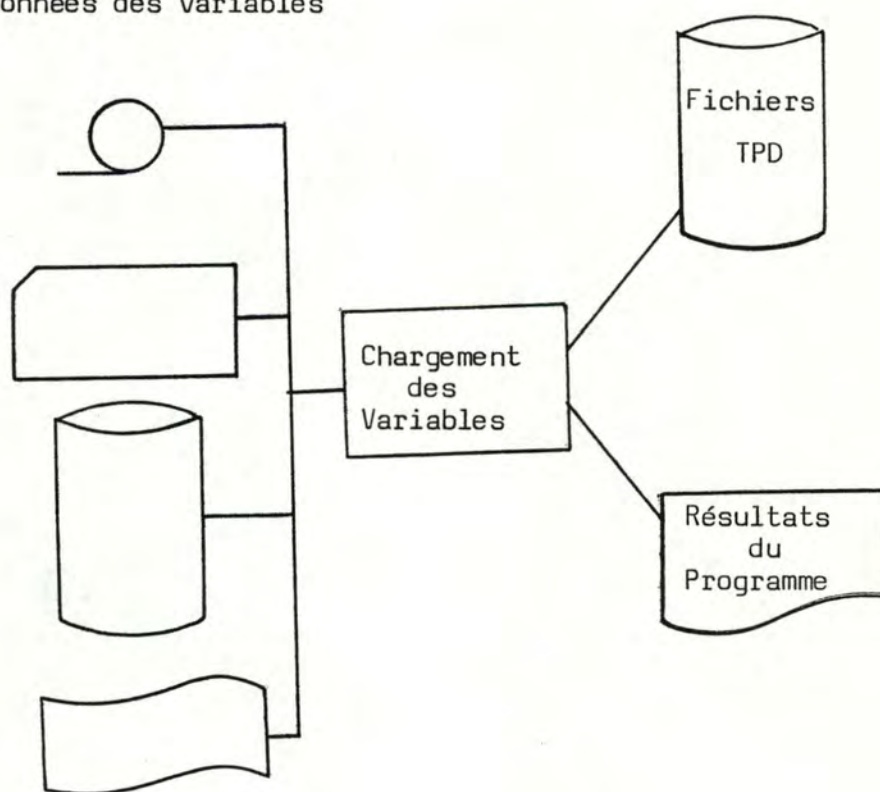
### a. Programmes généraux

Ils permettent la manipulation des données entre périphériques et leur transformation pour une meilleure présentation aux utilisateurs.

#### a.1. Programmes de chargement

Diagramme de flux :

Données des variables



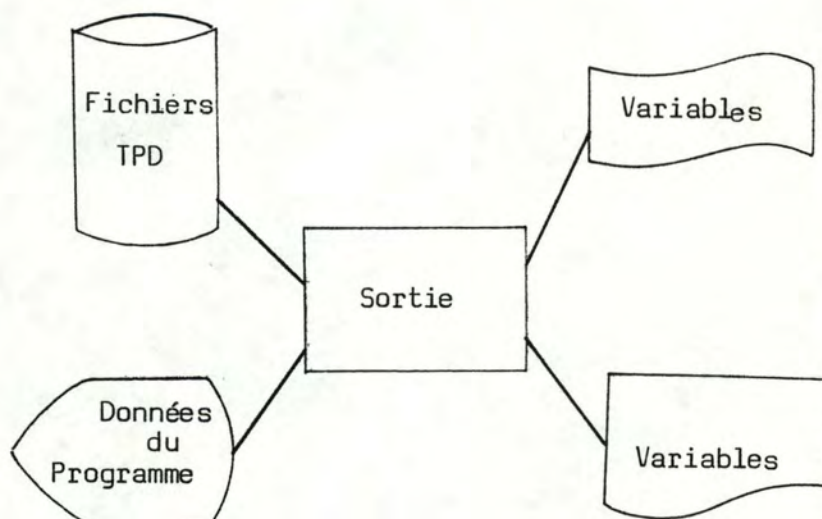
#### Description :

Les programmes de chargement permettent de charger sur disques magnétiques les données des variables à partir de bandes magnétiques, cartes perforées ou d'un fichier créé par l'utilisateur à l'aide d'un éditeur de texte.



## a. 2 Programmes de sortie

Diagramme de flux :



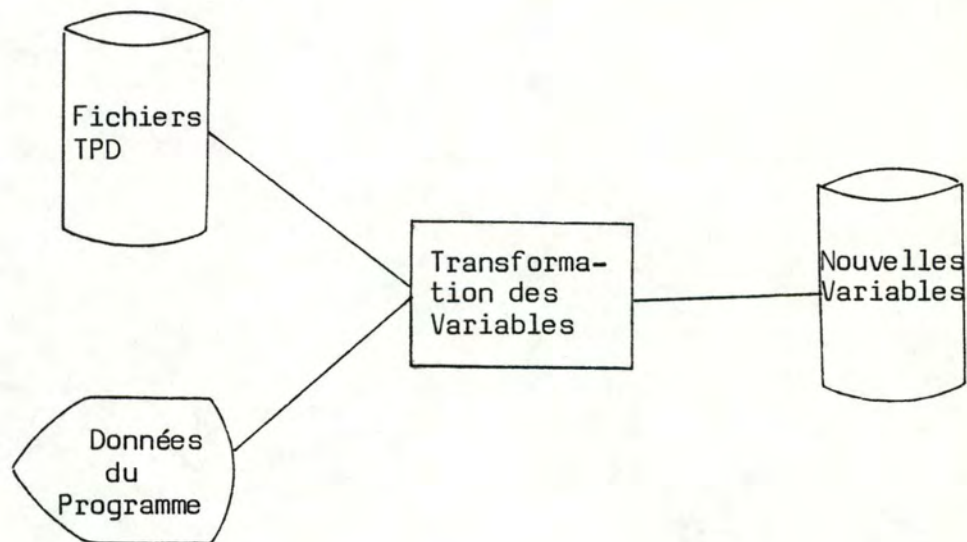
### Description :

Les programmes de sortie permettent au choix :

- d'imprimer les titres de toutes les variables enregistrées dans un fichier de titres,
- d'imprimer variable par variable les titres, paramètres et données des variables enregistrées dans un fichier TPD. A chaque impression des variables est associée l'impression de : moyenne, minimum, maximum, variance et écart type.

### a. 3 Programmes de transformation

Diagramme de flux :

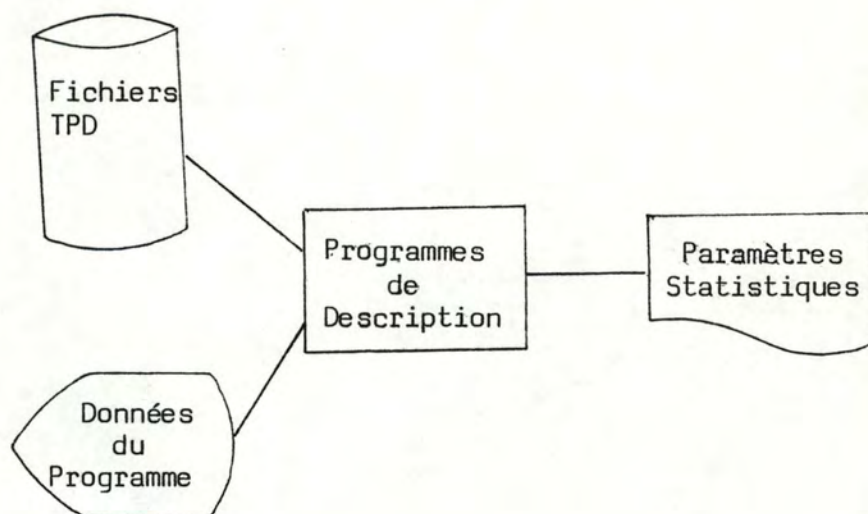


#### Description :

Les programmes de transformation permettent d'effectuer des transformations de variables et de calculer des fonctions de deux ou trois variables à partir des données enregistrées sur disque, les résultats étant eux-mêmes transférés sur disque comme de nouvelles variables.

#### a. 4 Programme de description

Diagramme de flux :



#### Description :

Les programmes de description sont destinés à établir des distributions de fréquence à une dimension (fréquences simples et fréquences cumulées) à les représenter graphiquement sous la forme d'histogrammes et à calculer les principaux paramètres (valeurs extrêmes, moyenne, écart type etc.).



b. Programmes de corrélation et de régression

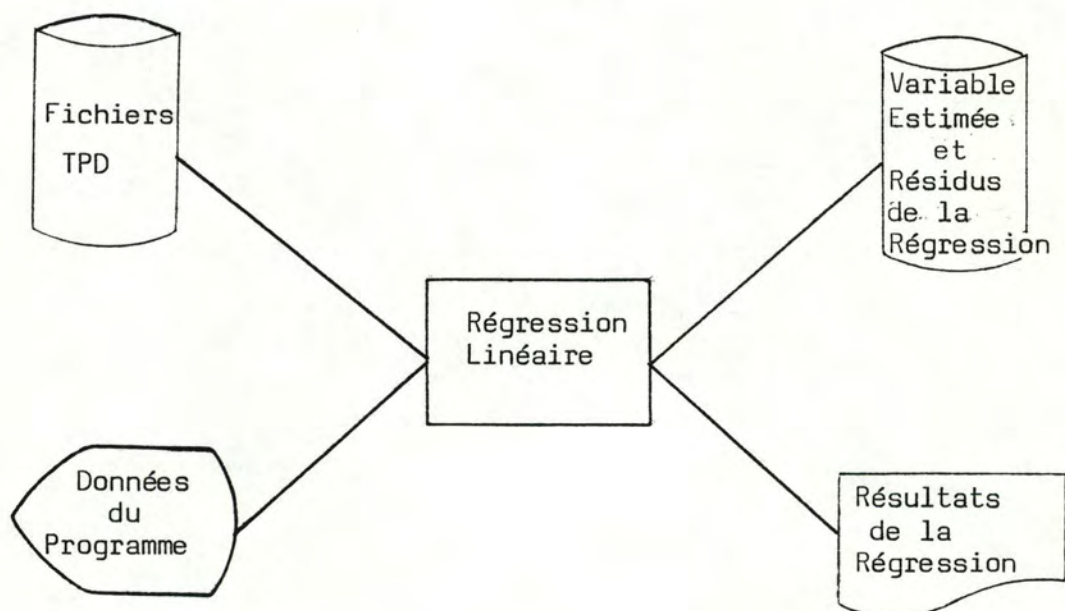
Ils englobent le calcul de matrices de corrélation, la régression linéaire et non linéaire simple et la régression multiple, l'établissement des fréquences à deux dimensions et les analyses de covariance à un et à deux critères de classification.

Ils sont divisés en :

- Programmes de calcul de matrices de corrélation.
- Programmes de régression linéaire simple.
- Programmes de régression multiple.

b. 1 Programme de calcul de matrices de corrélation

Diagramme de flux :

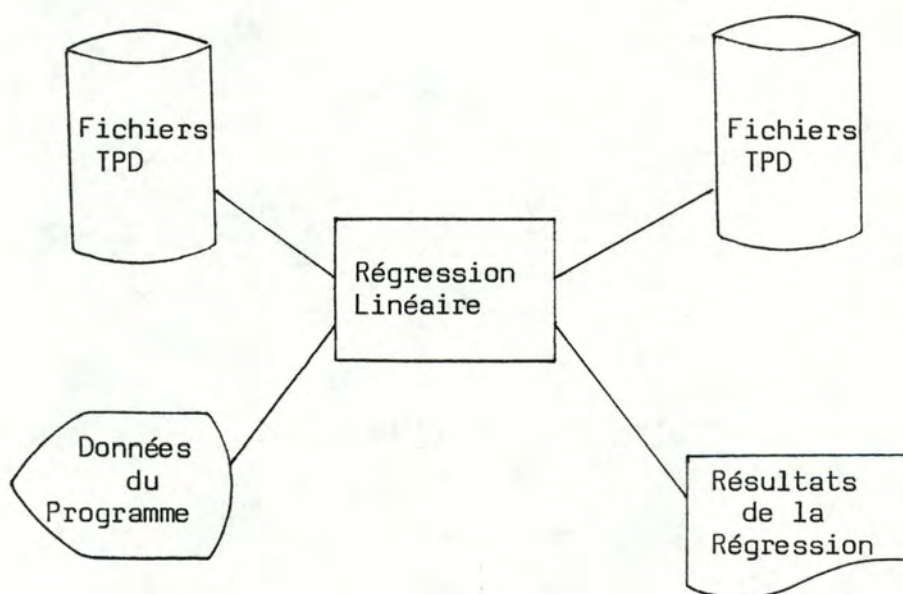


## Description :

Ce programme permet l'impression et l'enregistrement sur disque d'une matrice de corrélation.

b. 2 Programme de régression linéaire simple

## Diagramme de flux :

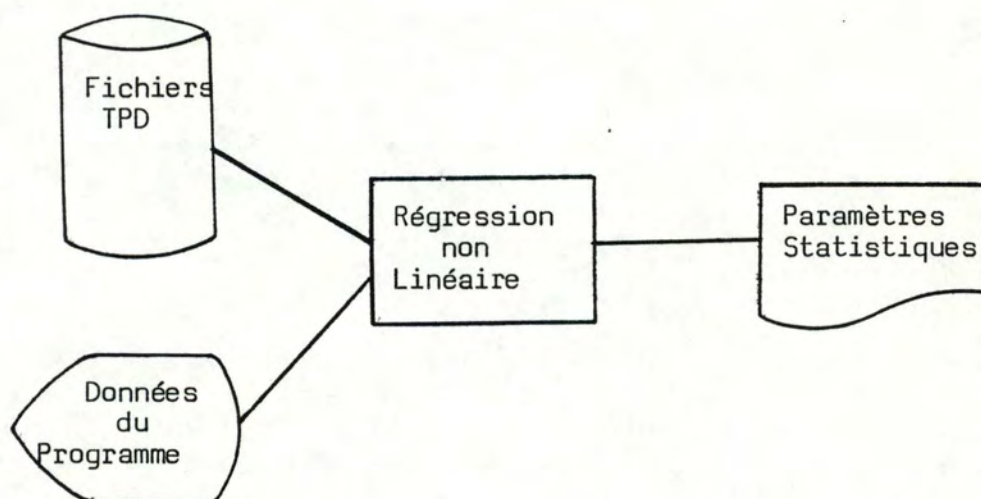


## Description :

Les programmes de régression linéaire sont destinés à établir des distributions de fréquence à deux dimensions et au calcul des principaux paramètres caractéristiques de séries statistiques : variance, écart type, coefficient de corrélation et de régression pour le modèle  $y = a + bx$ . De façon facultative on peut enregistrer sur disque les résidus de régression.

b. 3 Programmes de régression non linéaire simple

Diagramme de flux



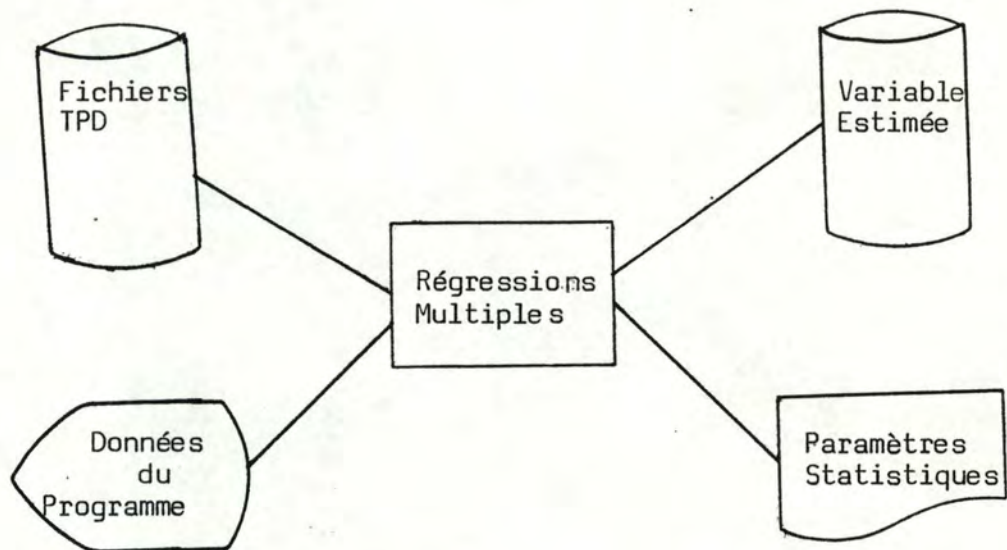
Description :

Les programmes de régression non linéaire permettent de résoudre tout problème de régression non linéaire quel que soit le modèle employé.



b. 4 Programme de régression multiple

Diagramme de flux



Description :

Les programmes de régression multiple calculent les paramètres de régression correspondants à un modèle linéaire :

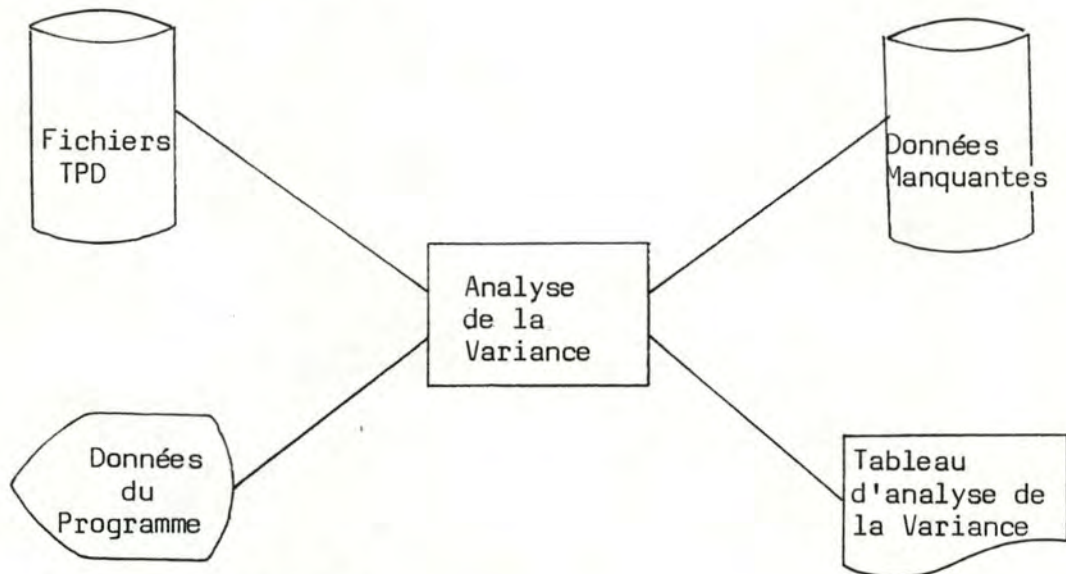
$$Y = b_0 + b_1 x_1 + b_2 x_2 + \dots + b_m x_m$$

Ils permettent d'écrire dans des fichiers TPD les valeurs de la variable dépendante estimées par l'équation de régression finale.

Ils calculent et impriment les paramètres statistiques des variables explicatives et dépendantes et les résultats partiels et finaux de la régression.

### c. Programmes d'analyse de la variance

Diagramme de flux :



Description :

Les programmes d'analyse de la variance permettent de traiter les cas d'analyse à un, deux, trois et quatre critères de classification pour différents modèles et, dans certains cas, la comparaison de moyennes.

L'estimation des données manquantes est faite dans les cas nécessaires. On imprime les valeurs extrêmes, les estimations de données manquantes, le tableau d'analyse de la variance et les résultats de tests particuliers.

### d. Programmes d'ajustement

Ils permettent de résoudre les problèmes d'ajustement de distributions théoriques et de traiter les tableaux de contingence à deux ou plusieurs dimensions.

## 2. 2. Structure actuelle des données

### 2.2.1. Dictionnaire de données

L'information de chaque variable considérée comporte quatre lignes de titres, dix paramètres et les données.

Entité VARIABLE

DEFINITION : Caractéristique étudiée sur un ou plusieurs individus.

IDENTIFIANT : num-var

ATTRIBUTS : - numéro de variable (abrevié num-var)

DEFINITION : Il identifie les variables parmi celles existantes

FORMAT : 1 à 3 chiffres entiers

VALEURS POSSIBLES : (1 à 500)

- titre

DEFINITION : Le titre contient des informations sur :

- Les expérimentateurs

- L'expérience considérée

- La variable

- Type de transformation sur la variable

Ex.  $X1 * 100$

Le contenu de chaque chaîne n'est pas standardisé et sert comme information générale associée à la variable.

FORMAT : 4 chaînes de 48 caractères

Entité PARAMETRES

DEFINITION : Donnent des indications sur les données, les nombres d'observation par échantillons et autres caractéristiques.



IDENTIFIANT : num-var

ATTRIBUTS : Varient selon l'analyse statistique à considérer.  
L'expérience a montré que pour les applications existantes 10 paramètres sont suffisants. Les paramètres utilisés dans l'analyse de la régression multiple sont :

P(6) : effectif des échantillons

P(8) : code de fin des données

P(9) : code de données manquantes

FORMAT : 10 Numéros réels

Entité DONNEES

DEFINITION : L'entité donnée est un ensemble de mesures d'une variable pour 1 ou plusieurs individus.

IDENTIFIANT : num-var

ATTRIBUTS : - num-var : (cf. entité variable)  
- enregistrement-données

DEFINITION : Ensemble de longueur variable des données entières

FORMAT : Enregistrement de longueur variable des numéros entiers

Remarque : Une donnée peut être :

- Un code de données manquantes.

DEFINITION : Il est une valeur donnée par l'utilisateur pour indiquer les mesures qui n'ont pas été réalisées.

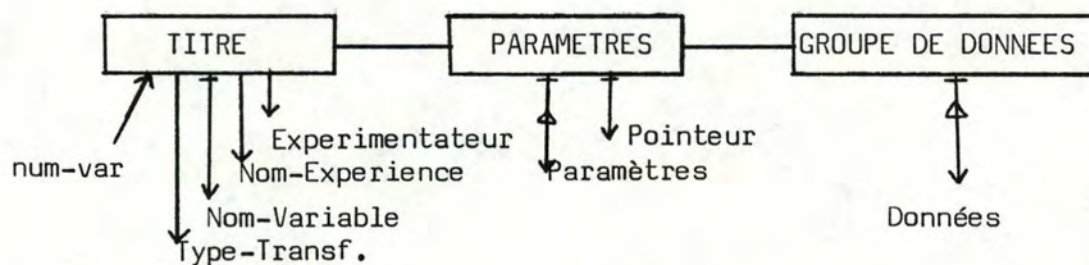
- Un code de fin des données.

DEFINITION : Il est une valeur donnée par l'utilisateur pour indiquer la fin d'un enregistrement des données.

VALEURS POSSIBLES : Le code de fin des données peut prendre différentes valeurs si l'on veut indiquer que les variables appartiennent à différents échantillons qui seront analysés ensemble.

### 2.2.2. Structure d'accès

La structure d'accès utilisée est la suivante :



### 2.2.3. Description des fichiers<sup>(1)</sup>

Les informations ont été stockées sur 3 fichiers à accès direct :

File F-TITRE contains TITRE

organisation is relative

record type TITRE

identifiant NUM-VAR

acces-key NUM-VAR

2 EXPERIMENTATEUR char(48)

2 NOM-EXPERIENCE char(48)

2 NOM-VARIABLE char(48)

2 TYPE-TRANSF char(48)

(1) La nomenclature utilisée est décrite dans HAINAUT (1984).



File F-PARAMETRE contains PARAMETRES

organisation is relative

record type paramètre

identifier NUM-VAR

acces-key NUM-VAR

2 PAR array (1..10) of reels

2 POINTEUR integer

File F-G DONNEES

organisation is relative

record type GROUPE-DE-DONNEES

acces key POINTEUR

2 GROUPE-DONNEES array (1..50) of integer

La fin des fichiers est signalée avec les caractères '\*\*' dans le dernier enregistrement du fichier des titres. Dans le fichier des paramètres, un pointeur signale le premier enregistrement libre du fichier des données.

Les fichiers de titres et paramètres ont été limités à 500 enregistrements et le fichier des données à 2000 enregistrements (50 données par enregistrement).

Tenant compte de l'organisation de ce matériel dans l'ordinateur actuel, il faut un mot de 24 bits pour stocker 4 caractères alpha-numériques ou un nombre entier et un groupe de deux mots pour un nombre réel. Cette dimension des fichiers implique l'occupation en mémoire magnétique de 134.500 mots, ce qui correspond environ à 400 K-octets (3).

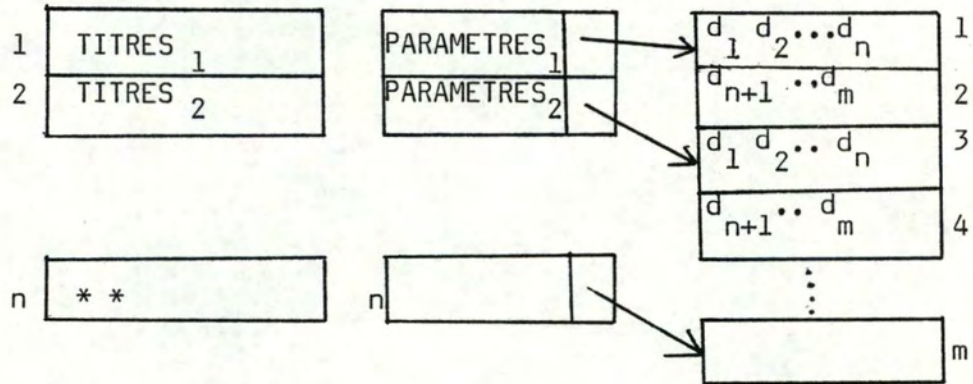
fichier T = (500) (4) (48) /4 = 24.000 mots

fichier P = (500) (10) (2) + (1) = 10.500 mots

fichier D = 100.000 mots



Graphiquement on peut représenter la structure de la façon suivante :



ou  $d_1 \dots d_m$  représentent les données.

### 2. 3 Critique de l'existant

Le système actuel présente des difficultés liées pour la plupart au traitement batch :

- Temps de réponse à l'utilisateur compris entre 1 heure et plusieurs jours dû au temps d'attente entre chaque traitement.
- Format d'entrée des données limité aux nombres entiers.
- Format rigide des sorties de résultats qui fait que certains utilisateurs reçoivent trop d'information. Ce cas se présente dans l'impression des paramètres statistiques à chaque pas de la régression.
- Perte de place en mémoire occupée par les données des variables non utilisées vu la difficulté de parvenir à leur élimination.
- Absence d'une description complète des algorithmes ce qui rend difficile la tâche de modifier les programmes pour les rendre interactifs.

### 2. 4 Expression des besoins

A court terme le centre de recherche disposera d'un ordinateur plus puissant qui permettra aux utilisateurs de travailler en temps réel; son utilisation devra permettre de résoudre les problèmes indiqués précédemment.

#### OBJECTIFS

Conception et implémentation d'un système interactif permettant :

- Une facile utilisation des programmes statistiques
- Bon temps de réponse
- Facile interprétation des résultats intermédiaires et finaux
- Adaptable à de petites machines

Le logiciel devra être portable, facilement extensible et si possible transparent au système d'exploitation en donnant des résultats fiables.



## 2. 5 Ebauche de solutions

Le passage de mode batch à interactif doit permettre d'atteindre les objectifs généraux.

On envisage trois possibilités :

- Utiliser le système actuel en ajoutant des messages interactifs pour l'utilisateur.
- Elaborer un système qui reprend les analyses les plus utilisées de façon cohérente et coordonnée.
- Elaborer des programmes interactifs indépendants qui auraient accès à une base de données communes.

La première solution conduit à garder un système dépendant de la structure de données et difficilement modifiable du au fait qu'il est écrit en FORTRAN IV, langage qui ne possède pas d'instructions permettant un style de programmation structuré.

La deuxième solution requiert une analyse complète qui ne peut pas être faite à court terme, étant donné la complexité des traitements à effectuer.

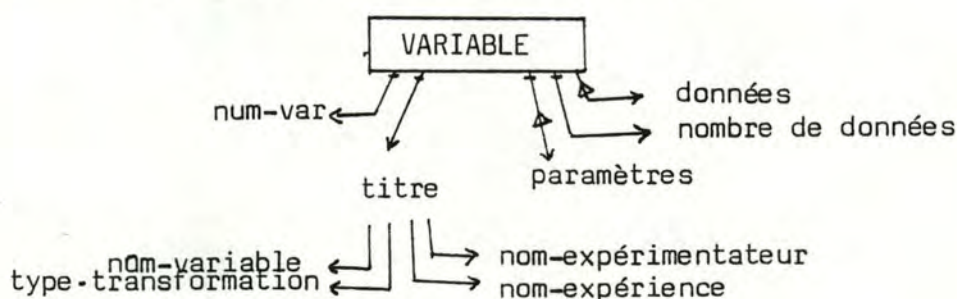
La troisième solution permet d'atteindre les objectifs généraux et peut être implémentée progressivement en utilisant le compilateur FORTRAN 77 qui permet de programmer de façon structurée.

## 2. 6 Choix d'une structure de données

On a analysé trois alternatives de structuration de fichiers :

- a) Créer un fichier séquentiel avec des articles de longueurs variables.

Représentation du type d'article :





Description du fichier :

file F-VARIABLE

organisation is sequentiel

record type VARIABLE

2 NUM-VAR integer (4)

2 TITRE

3 NOM-EXPERIMENTATEUR char (48)

3 NOM-EXPERIENCE char (48)

3 NOM-VARIABLE char (48)

3 TYPE-TRANSFORMATION char (48)

2 PARAMETRES array (1..10) of reels

2 NDONNEES integer (4)

2 DONNEES array (1.. NDONNEES) of reels

Avantages :

Facilité de programmation et maintenance

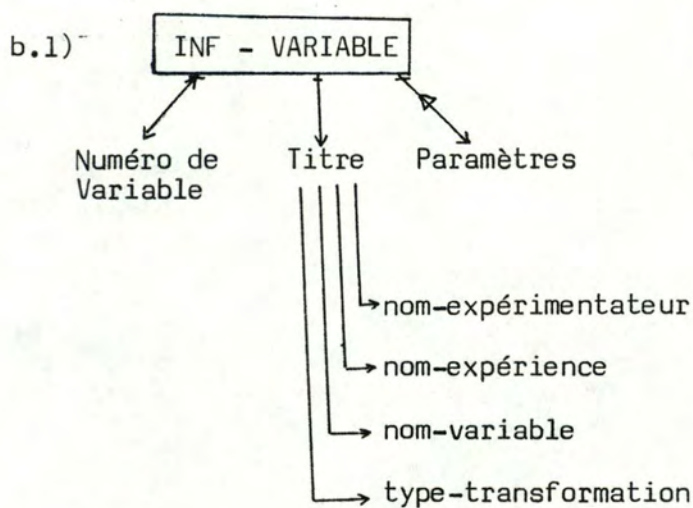
Bonne utilisation du support

Inconvénients :

Coûteuse en temps de traitement en comparaison avec des techniques d'accès directs pour cette application, car la plupart des accès sont directs sur base de numéro de variable.

- b) Créer 2 fichiers séquentiels indexés, 1 fichier pour l'information fixe et l'autre pour l'information variable.

Représentation du type d'article :



Description du fichier :

file F-INF-VARIABLE

organisation is indexed; access key is NUM-VAR

record type INF-VARIABLE

2 NUM-VAR integer (4)

2 TITRE

3 NOM-EXPERIMENTATEUR char(48)

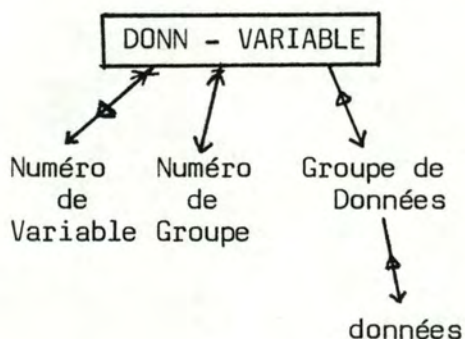
3 NOM-EXPERIENCE char(48)

3 NOM-VARIABLE char(48)

3 TYPE-TRANSF char(48)

2 PAR array (1..10) of reels

b.2) Représentation du type d'article :



Description du fichier :

file F-DONN-VARIABLE

organisation is indexed, acces key are NUM-VAR, NUM-GROUPE

record type DONN-VARIABLE

2 NUM-VAR integer (4)

2 NUM-GROUPE integer (5)

2 NUM-GROUPE-DONNEES array (1..50) of reels

Avantages : Cette organisation facilite la programmation et la maintenance si on dispose d'un outil pour gérer les fichiers séquentiels indexés; elle est en plus conseillée car le nombre d'accès aléatoires est élevé par rapport au nombre d'accès séquentiels.



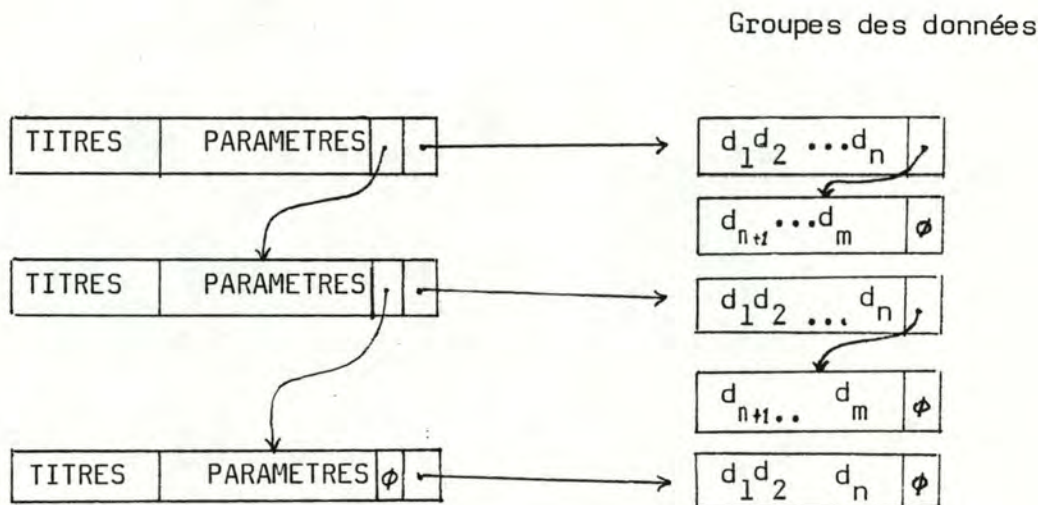
- c. Créer un fichier multiliste avec la structure d'accès décrite dans 2.2.2. (avec la différence dans l'implémentation)

Les fichiers des titres et paramètres à accès direct permettent l'implémentation d'une liste enchaînée avec des pointeurs gérés de façon indépendante du fichier de base. A chaque titre est associé en plus un pointeur vers les enregistrements des données.

Deux listes en forme de pile donnent la position relative des enregistrements des titres et des données disponibles.

A chaque opération d'ajoute ou d'élimination d'une variable, on actualise les piles des enregistrements disponibles.

Graphiquement on peut représenter la structure de la façon suivante :



ou  $d_1 \dots d_m$  représentent les données

**Avantages :** Les opérations d'ajoute et d'élimination n'impliquent pas le parcours de tout le fichier comme dans le cas de fichiers séquentiels ou dans la structure multiliste du paragraphe 1.2.2. Les enchaînements sont fait par des actualisations de trois listes de pointeurs implémentés indépendamment des fichiers de base.

Cette structure à été choisie pour les avantages cités.

## Chapitre 3 : ANALYSE DE LA REGRESSION MULTIPLE

### 3. 1. Introduction

La régression multiple a pour but l'étude de la relation qui peut exister entre une variable observée dite variable dépendante et deux ou plusieurs autres variables dites indépendantes ou explicatives. On fait l'analyse à partir de  $n$  observations de la variable dépendante correspondant chacune à  $p$  valeurs des variables explicatives. On peut représenter une observation de la façon suivante :

$$\underline{Y}_i = f( X_{i1} \dots X_{ip} ) + \underline{R}_i ; i=1, n \quad (3. 1)$$

ou :

$\underline{Y}_i$  = variable dépendante

$\underline{R}_i$  = variable aléatoire appelée résidu

$X_{i1} \dots X_{ip}$  : variables explicatives

La relation 3. 1 et les hypothèses faites sur  $\underline{R}_i$  constituent le modèle de régression. Généralement on suppose que le résidu est d'espérance nulle de sorte que  $f(X_{i1} \dots X_{ip})$  est la valeur attendue de la variable  $\underline{Y}_i$ . Dans la plupart des problèmes réels la fonction  $f$  dépend de paramètres inconnus. Les problèmes de régression demandent la solution des deux problèmes :

- Il faut choisir le modèle c'est-à-dire la nature et le contenu de l'équation à utiliser.
- Estimer les paramètres qui interviennent dans ce modèle. Parmi les différents modèles de régression, le cas où la fonction  $f$  est linéaire en ses paramètres est le plus facile à traiter.



### 3.2. Terminologie et définitions (5).

Les principaux paramètres statistiques utilisés dans le calcul et l'analyse de la régression multiple sont :

- Somme des carrés des écarts de la variable  $j$  :  $SCE_j$

$$SCE_j = \sum_{\alpha=1}^n (x_{\alpha j} - \bar{x}_j)^2 = \sum_{i=1}^n x_{ij}^2 - \frac{1}{n} \left( \sum_{i=1}^n x_{ij} \right)^2 \quad (3.2)$$

- Somme des produits des écarts  $SPE_{ij}$

$$SPE_{ij} = \sum_{\alpha=1}^n (x_{\alpha j} - \bar{x}_j) (x_{\alpha i} - \bar{x}_i) \quad (3.3)$$

$$= \sum_{\alpha=1}^n x_{\alpha j} x_{\alpha i} - \frac{1}{n} \left( \sum_{\alpha=1}^n x_{\alpha i} \right) \left( \sum_{\alpha=1}^n x_{\alpha j} \right)$$

- Somme des carrés des écarts résiduelle SCERES

$$SCERES = \sum_{i=1}^n [ \underline{y}_i - \underline{y}(x_i) ]^2 \quad (3.4)$$

$\underline{y}_i$  = valeurs observées

$\underline{y}(x_i)$  = valeurs estimées par une équation de régression

- Variance résiduelle de  $y$  : est la variance des résidus de la régression, elle est estimée par :

$$\hat{\sigma}_{y.x}^2 = \frac{1}{n-p-1} \sum_{i=1}^n [ \underline{y}_i - \underline{y}(x_i) ]^2 = \frac{1}{n-p-1} SCERES \quad (3.5)$$

ou



- Degrés de liberté résiduels

$$DLRES = n - p - 1 \quad (3.6)$$

$n$  = nombre d'observations

$p$  = nombre de variables dans l'équation de régression

- Carré moyen résiduel

$$CMRES = SCERES / DLRES \quad (3.7)$$

- Degrés de liberté de la régression

$$DLRES = p$$

- Somme des carrés des écarts de la régression

$$\begin{aligned} SCEREG &= \sum ( \underline{y}(X_i) - \bar{y} )^2 \\ &= SCE_y - SCERES \end{aligned} \quad (3.8)$$

- Carré moyen de la régression

$$CMREG = SCEREG / p \quad (3.9)$$

$\chi^2$  Les variables aléatoires CMRES et CMREG suivent une distribution et la relation.

$$F = CMREG / CMRES \quad (3.10)$$

Suit une distribution F de Snedecor avec  $p$  et  $n-p-1$  degrés de liberté.

Ecart type résiduel ECTRES

$$ECTRES = (CMRES)^{0.5} \quad (3.11)$$

- Covariance :  $\text{Cov} (X,Y)$

$$\text{Cov} (X,Y) = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{X}) (y_i - \bar{Y}) \quad (3.12)$$

La covariance caractérise simultanément deux séries d'observations. Elle est positive ou négative selon que la relation entre les deux séries de données est croissante ou décroissante.

- Coefficient de corrélation entre X et Y :

Est une mesure de l'association entre les variables aléatoires  $\underline{X}$  et  $\underline{Y}$ .

$$C_{xy} = \frac{\text{Cov} (X,Y)}{s_x s_y} \quad (3.13)$$

$s_x, s_y$  = écarts types des variables X et Y

- Coefficient de corrélation partielle .

Mesure l'intensité de la relation entre deux variables indépendamment de l'influence d'une troisième variable.

Ex. dans le cas de trois variables, le coefficient de corrélation partielle entre Y et Z est le coefficient de corrélation entre les résidus  $Y - Y(X)$  et  $Z - Z(X)$  des régressions linéaires à deux dimensions.

$$C_{yz.x} = \frac{C_{yz} - C_{xy} C_{xz}}{(1 - C_{xy}^2)(1 - C_{xz}^2)}$$

Dans le cas général de plusieurs variables, le coefficient de corrélation partielle de  $X_p$  et  $Y$  par exemple peut être défini comme étant le coefficient de corrélation des résidus de la régression linéaire de  $Y$  en fonction de  $X_1, \dots, X_{p-1}$  avec les résidus de la régression linéaire de  $X_p$  en fonction de  $X_1, \dots, X_{p-1}$ .

Il peut être calculé par la formule :

$$C_p = \frac{\hat{a}_{dd}^{(*)}}{(\hat{a}_{ii} \hat{a}_{dd})^{1/2}} \quad (3.14)$$

---

(\*)  $\hat{a}$  est définie dans le paragraphe 3.4.1



- Coefficient de corrélation multiple R

Pour une variable dépendante  $y$  et  $p$  variables explicatives  $X_1, X_2 \dots X_p$  le coefficient de corrélation multiple  $R$  est le coefficient de corrélation des valeurs observées  $Y_i$  de la variable dépendante avec les valeurs estimées  $\hat{Y} (X_i)$  obtenues à l'aide de l'équation de régression multiple.

Ce coefficient peut être calculé par :

$$R = SC_{REG} / SCE_y \quad (3.15)$$

- Coefficient de détermination =  $R^2$

Estimation du coefficient de corrélation multiple et du coefficient de détermination multiple.

Les valeurs observées  $R$  des coefficients de corrélation multiple sont utilisées comme estimations  $P$  des valeurs théoriques, mais les estimations ainsi obtenues ne sont pas absolument correctes.

Pour le coefficient de détermination, l'estimation

$$P^2 = 1 - \frac{n-1}{n-p-1} (1 - R^2) \quad (3.16)$$

est plus exacte que l'estimation directe :  $P^2 = R^2$  (5,335).

### 3. 3. Modèle et Hypothèse

Dans ce travail on va considérer le modèle linéaire suivant :

$$Y_1 = b_0 + b_1 X_{11} \dots \dots \dots + b_p X_{1p} + R_1 \quad (3. 17)$$

$$Y_n = b_0 + b_1 X_{n1} \dots \dots \dots + b_p X_{np} + R_n$$

En notation matricielle on a :

$$Y = XB + R \quad (3. 18)$$

où les valeurs de la variable  $X_{i0}$ ,  $i=1, n$  sont égaux à l'unité.

Si on applique le modèle linéaire dans le cas où on souhaite l'étude d'une variable  $Y$  en fonction de  $p$  variables explicatives pour lesquelles on dispose de  $n$  observations, on a :

$Y$  : Vecteur aléatoire  $n \times 1$  observable

$X$  : Matrice  $(n \times p)$  des valeurs connues  $(p+1) \times 1$

$B$  : Vecteur des paramètres inconnus

$R$  : Vecteur aléatoire  $(n \times 1)$  non observable

### 3. 3. 1. Les conditions d'application

Ce modèle peut-être appliqué dans les cas où :

- Les variables explicatives sont des variables non aléatoires qui ont des valeurs connues sans erreur.
- Les résidus ont tous une même distribution normale de moyenne nulle, variance constante et doivent être indépendants les uns des autres ce qui implique que les valeurs de la variable dépendante sont extraites de distributions normales indépendantes et de même variance.
- Le rang de  $X$  est égal à  $p$  ce qui implique que aucune colonne de  $X$  n'est combinaison linéaire des autres.
- Le modèle linéaire peut-être appliqué également dans le cas où les  $p$  variables explicatives ou certaines d'entre elles sont des variables aléatoires dont les valeurs sont observées dans des conditions analogues à celles de la variable dépendante.

Des informations complémentaires peuvent être trouvées dans DAGNELIE (1982) et DRAPER et SMITH (1966).



### 3. 4. Estimation des paramètres

Il s'agit de trouver B et R dans la relation (3. 18) à partir des valeurs de Y observées et des valeurs de X.

Si on considère  $E(R) = 0$ , on peut écrire le modèle de la façon suivante :

$$E(Y) = X B \quad (3.19)$$

La méthode la plus utilisée est celle des moindres carrées, et consiste à choisir comme estimateur de B le vecteur  $\hat{B}$  qui minimise la somme des carrées des résidus  $h(B)$ .

$$h(B) = r_1^2 + r_2^2 + r_3^2 + \dots + r_n^2 = R'R$$

$$= (Y - XB)' (Y - XB)$$

$$= Y'Y - 2Y'XB + B'X'XB$$

$$\frac{d h(B)}{d B} = -2X'Y + 2 X' XB \quad (3.10)$$

En annulant la dérivée, on obtient l'estimateur  $\hat{B}$  de B :

$$X'XB = X'Y \quad (3.21)$$

Si le modèle est de rang maximum le système d'équations (3.21) admet une solution unique :

$$\hat{B} = (X'X)^{-1} X'Y \quad (3.22)$$

$$\hat{Y} = X \hat{B}$$

Si au lieu de X on utilise une matrice centrée XC où ses éléments ont été obtenus par :

$$XC_{ij} = X_{ij} - XM_{.j}$$

où :

$$XM_{.j} = (1/n) \sum_{i=1}^n X_{ij} ;$$

et si au lieu de Y, on utilise un vecteur centré YC, où

$$YC_i = Y_i - YM ;$$

$$YM = (1/n) \sum_{i=1}^n Y_i$$

On peut exprimer (3.22) en termes de sommes de carrés et de produits des écarts :

$A_{xx}$  : matrice des sommes des carrés et des produits des écarts des variables explicatives.

$$= XC'XC$$

$a_{xy}$  : vecteur-colonne des sommes de produits des écarts des variables explicatives et de la variable dépendante.

$$= XC'YC$$

$a_{yy}$  : somme des carrés des écarts de la variable dépendante.

En remplaçant dans l'équation (3.21) on a :

$$A_{xx} \hat{B} = a_{xy} \quad (3.22)$$

dont la solution est :

$$\hat{B} = (A_{xx})^{-1} a_{xy} \quad (3.23)$$

l'équation de régression multiple est donc :

$$Y = \bar{Y} + B(X - \bar{X}) \quad (3.24)$$

On peut donc obtenir une estimation du terme indépendant par :

$$= \hat{b}_0 = \bar{Y} - B\bar{X} \quad (3.25)$$

Des informations complémentaires relatives à l'obtention des formules 3.24 et 3.25 se trouvent dans DRAPER et SMITH (1966).

Les différentes estimations ne sont définies que quand la matrice  $A_{xx}$  est non singulière, c'est-à-dire quand toutes les variables explicatives sont linéairement indépendantes. Lorsque, par contre, la matrice  $A_{xx}$  est singulière, les coefficients de régression partielle sont indéterminés.

Les distributions d'échantillonnage des coefficients de régression  $b_i^{(n)}$  sont liés aux distributions  $t$  de Student. Leurs erreurs standards  $S_1$  et leurs limites de confiance sont respectivement :

$$\sqrt{\hat{\sigma}_{y.x}^2 a_{ii}} \text{ et } b_i \pm t_{1-\alpha/2} \sqrt{\hat{\sigma}_{y.x}^2 a_{ii}} \quad (3.26)$$

Le symbole  $a_{ii}$  désignant le terme de la matrice inverse  $A_{xx}$  et le nombre de degrés de liberté de la variable  $t$  étant égal à  $n-p-1$ .

---

(1)  $b_i$  représente l'occurrence  $i$  du vecteur  $B$ .



### 3. 4. 1. Techniques de calcul

Des calculs basés sur l'équation 3.7 ne présentent pas de difficultés pour deux ou trois variables indépendantes. Les résultats avec plusieurs variables indépendantes et beaucoup de données peuvent être complètement invalidés à cause des erreurs d'arrondi qui donnent une mauvaise stabilité numérique. Pour cette raison ces algorithmes sont considérés comme les plus mauvais (SCHIFLERS 1.978).

Il existe des méthodes permettant d'ajouter ou de supprimer une variable explicative dans une équation de régression, sans que l'ensemble des calculs soit complètement repris.

Exemple : dans le cas de suppression d'une variable explicative dans une équation à  $p$  variables, il est possible de calculer les  $p - 1$  coefficients de régression partielle et la variance résiduelle à partir des éléments correspondants à l'équation initiale de  $p$  variables.

Une méthode couramment employée pour l'estimation des coefficients de régression est présentée par Jennrich (1978) est basée sur l'application d'un opérateur appelé "sweep" qui permet d'effectuer des transformations linéaires sur les lignes et colonnes d'une matrice.

Le résultat de l'application de cet opérateur à une matrice  $A$  en utilisant l'élément  $k$ -ème de la diagonale, s'il est différent de zéro, est une nouvelle matrice  $\hat{A}$  dont les éléments sont :

$$\begin{aligned}\hat{a}_{kk} &= -1 / a_{kk} \\ \hat{a}_{ik} &= a_{ik} / a_{kk} \\ \hat{a}_{kj} &= a_{kj} / a_{kk} \\ \hat{a}_{ij} &= a_{ij} - (a_{ik} \times a_{kj}) / a_{kk}\end{aligned}\tag{3.27}$$

pour  $i \neq j$  et  $j \neq k$

Une opération "sweep" peut être annulée avec les transformations linéaires suivantes :

$$\begin{aligned}
 \hat{a}_{kk} &= -1 / a_{kk} \\
 \hat{a}_{ik} &= -a_{ik} / a_{kk} \\
 \hat{a}_{kj} &= -a_{kj} / a_{kk} \\
 \hat{a}_{ij} &= a_{ij} - (a_{ik} \times a_{kj}) / a_{kk}
 \end{aligned} \tag{3.28}$$

pour  $i \neq k$  et  $j \neq k$

Jenrich énonce et fait la démonstration du théorème suivant :

Si on a une matrice partitionnée :

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \tag{3.29}$$

et s'il est possible d'appliquer l'opérateur sweep à la matrice A pour chaque élément de la diagonale de la matrice carrée  $A_{11}$ , on obtient comme résultat la matrice :

$$\hat{A} = \begin{bmatrix} -A_{21}^{-1} & A_{11}^{-1} A_{12} \\ A_{21} A_{11}^{-1} & A_{22} - A_{21} A_{11}^{-1} A_{12} \end{bmatrix} \tag{3.30}$$

En remplaçant  $A_{11}$ ,  $A_{12}$ ,  $A_{21}$ , et  $A_{22}$  par les matrices de sommes de carrés et de produits des écarts on a :

$$A = \begin{vmatrix} A_{xx} & a_{xy} \\ a'_{xy} & a_{yy} \end{vmatrix} \quad (3.31)$$

$$\hat{A} = \begin{vmatrix} -A_{xx}^{-1} & A_{xx}^{-1} a_{xy} \\ a'_{xy} A_{xx}^{-1} & A_{yy} - a'_{xy} A_{xx}^{-1} a_{xy} \end{vmatrix} \quad (3.32)$$

En tenant compte de l'équation (3.23), on obtient par cette méthode une estimation des coefficients de régression  $B$ .

$$\hat{A} = \begin{vmatrix} -A_{xx}^{-1} & \hat{B} \\ \hat{B}' & A_{yy} - a'_{xy} A_{xx}^{-1} a_{xy} \end{vmatrix} \quad (3.33)$$

La relation (3.27) est utilisée pour ajouter la variable  $X_k$  à l'équation de régression; dans ce cas l'élément  $\hat{a}_{ii}$  de la diagonale de la matrice  $\hat{A}$  est négatif. La relation 3.28 est utilisée quant il s'agit d'éliminer la variable  $X_k$  de l'équation de régression.



### 3. 5. Choix des variables explicatives

Supposons qu'on souhaite établir une équation de régression de  $Y$  en terme des variables indépendantes  $X_1, X_2 \dots X_p$ . Deux critères s'opposent pour sélectionner une équation résultante :

1. Faire l'équation utile pour des prédictions et pour cette raison inclure dans le modèle la plupart des variables indépendantes de telle façon qu'on puisse obtenir un bon ajustement.
2. Inclure la plus petite quantité possible des variables indépendantes vu la difficulté d'obtenir l'information pour une grande quantité de variables.

La sélection de la meilleure équation de régression cherche un compromis entre les deux critères. Dans certains cas aucun choix ne doit être fait car on souhaite que toutes les variables observées soient prises en considération. Tel cas peut se présenter lorsque l'objectif poursuivi est plus d'expliquer un phénomène que d'essayer de le prévoir.

Il existe plusieurs procédures statistiques qui essaient de résoudre ces problèmes; toutes sont largement utilisées, mais elles ne donnent pas toujours les mêmes résultats pour un problème donné. Toutes essaient d'assurer une précision maximum de l'équation de régression, c'est-à-dire une variance résiduelle minimum, en donnant la préférence à des variables explicatives fortement corrélées avec la variable dépendante et faiblement corrélées entre elles. On peut citer les procédures suivantes :

- a. Comparaison de toutes les équations de régression possibles.
- b. Sélection régressive.
- c. Sélection progressive des variables.
- d. Sélection mixte ou régression pas à pas.

### 3.5.1 Comparaison de toutes les équations de régression possibles

Différentes méthodes de calcul ont été proposées pour comparer toutes les équations de régression possibles. Si on a  $p$  variables explicatives on aura besoin de :

$p$  équation de régression simple

$p(p-1)/2$  équations de régression à deux variables explicatives, etc.

Cette façon de procéder est peu utilisée car elle demande beaucoup de calculs. Pour  $p$  variables explicatives le nombre total d'équations différentes est :

$$2^p - 1$$

### 3.5.2 Sélection régressive

Cette méthode est une amélioration de la méthode de comparaison de toutes les équations. Le choix des variables se réalise de la façon suivante :

1. Calculer une équation de régression qui contienne toutes les variables.
2. Pour chaque variable traitée comme si elle était la dernière variable à entrer dans l'équation de régression, effectuer l'analyse de la variance pour tester si la diminution de la variance résiduelle est significative.

La plus petite valeur du test  $F$  ex.  $F_p$  est comparée avec un niveau de signification  $F_o$  présélectionné.

Si  $F_p > F_o$  prendre l'équation de régression comme elle a été calculée. Avec cette méthode, à chaque pas, on élimine la variable dont la contribution est la moins importante jusqu'au moment où tous les coefficients de régression partiels sont significatifs par rapport à une valeur prédéterminée.



Cette méthode demande moins de calculs que la comparaison de toutes les équations et présente l'avantage de fournir des indications relatives à toutes les variables explicatives car toutes interviennent dans la première équation prise en considération; cependant elle ne conduit pas nécessairement à la solution optimale et si la matrice  $A_{xx}$  est presque singulière, cette procédure n'a pas de sens pour les erreurs d'arrondi.

### 3.5.3 Sélection progressive

Dans cette méthode on choisit les variables de proche en proche en assurant chaque fois le minimum de la somme des carrés des écarts ou de la variance résiduelle.

L'ordre d'insertion est déterminé en utilisant les coefficients de corrélation partielles comme une mesure de l'importance des variables qui ne se trouvent pas encore dans l'équation. On procède de la façon suivante :

- Choisir la variable explicative la plus fortement corrélée avec la variable dépendante.
- En deuxième lieu ajouter la variable qui avec celle déjà choisie assure un minimum de la variance résiduelle.
- Continuer ainsi jusqu'au moment où l'introduction d'une nouvelle variable ne provoque plus de réduction significative de la variance résiduelle.

Pour chaque variable à ajouter à l'équation de régression on examine :

- Le coefficient de corrélation multiple.
- Le test F pour la variable récemment ajoutée, ce qui montre si la variable a produit une diminution significative de la somme des carrés des écarts résiduelle en comparaison à celle produite par les variables qui se trouvaient dans l'équation de régression.

Au moment où la valeur F correspondant à la variable la plus récemment entrée devient non significative, le processus est terminé.



Cette procédure a l'avantage de demander moins de calculs car on n'examine pas plus de variables que celles qui sont nécessaires à chaque pas.

Elle présente l'inconvénient de ne pas examiner l'effet que l'introduction d'une nouvelle variable peut avoir sur le rôle joué pour une variable qui était entrée auparavant.

La méthode de sélection progressive est une des plus utilisées et permet d'obtenir des résultats intermédiaires qui donnent une information statistique valable à chaque pas de calcul. Il est possible d'obtenir des équations de régression intermédiaires et équation de régression multiple complète à la fin. Les équations intermédiaires sont construites en ajoutant une variable à la fois, variable qui soit statistiquement significative.

#### 3.5.4 Méthode de sélection mixte ou de régression pas à pas

Cette méthode est une amélioration de la procédure de régression progressive. On procède à l'introduction successive des variables mais à chaque pas de la régression on examine la signification des variables explicatives déjà présentes dans l'équation.

Cette méthode a l'avantage d'écarter de l'équation les variables nuisibles. En effet, une variable peut être significative au début de l'algorithme et ne plus l'être après l'addition de nouvelles variables. Dans ce cas la variable sera éliminée de l'équation de régression avant d'ajouter une variable additionnelle.

On introduit les variables explicatives jusqu'au moment où l'introduction d'une nouvelle variable ne provoque plus de réduction significative de la variance résiduelle. La méthode de régression pas à pas est considérée comme la meilleure parmi les méthodes de choix des variables, mais elle ne donne pas toujours la solution optimale. C'est la raison pour laquelle il faut un bon jugement pour choisir les variables initiales.

Des algorithmes présents dans BIOMEDICAL COMPUTER PROGRAMS (1971) prennent en considération ce problème en donnant à l'utilisateur la possibilité de décider que certaines variables soient obligatoirement présentes dans l'équation de régression.

### 3. 6. Test de signification des coefficients de régression

Les distributions d'échantillonnage des coefficients de régression sont liées aux distributions t de Student avec n-p-1 degrés de liberté. La variance de la distribution d'échantillonnage de  $b_i$  est donnée par :

$$\frac{\hat{\sigma}_{y.x}}{SCE_{x_i}} \quad (3.34)$$

et leur erreur standard, par :

$$s_i = \hat{\sigma}_{y.x} / \sqrt{SCE_{x_i}}$$

Le test de signification d'un coefficient de régression  $b_i$  peut être réalisé en calculant la valeur :

$$t_{obs} = b_i / s_i \quad (3.35)$$

où  $b_i$  = valeur estimée du coefficient de régression, et on rejette l'hypothèse nulle

$$H_0 : b_i = 0 \quad (3.36)$$

au niveau  $\alpha$  lorsque

$$t_{obs} \geq t_{1-\alpha/2}$$

avec n-p-1 degrés de liberté.



D'où une variable dans l'équation de régression sera significative si son coefficient de régression est significatif.

Le test de signification des coefficients de régression partielle peut être également réalisé par une analyse de la variance.

L'hypothèse à tester est la suivante :

$H_0$  = La diminution de la variance résiduelle due à l'introduction de la variable  $X_i$  est nulle

La valeur de  $F_c$  de Snedecor pour tester cette hypothèse est :

$$F_c = \frac{SCE_{\text{due à } X_i / X_1 \dots X_{p-1}}^{(*)} / 1}{SCE_{\dots 1 \dots i \dots p} / n-p-1} \quad (3.37)$$

$$SCE_{\text{due à } X_i / X_1 \dots X_p} = SCE_{Y, X_1 \dots X_{p-1}} - \frac{SCE_{Y, X_1 \dots X_i \dots X_p}}{SCE_{Y, X_1 \dots X_i \dots X_p}}$$

On rejette l'hypothèse nulle à un niveau de signification  $\alpha$  si

$$F_c > F_{1-\alpha/2}$$

La valeur  $F$  est utilisée dans la méthode de régression pas à pas comme valeur limite pour introduire et éliminer une variable de l'équation de régression. Elle sera présentée respectivement comme  $F_{\text{entrée}}$  et  $F_{\text{sortie}}$ .

---

(\*)  $SCE_{\text{due à } X_i / X_1 \dots X_{p-1}}$  que nous noterons  $SCE_{X_i / X_1 \dots X_{p-1}}$  :  
 somme des carrés des écarts liée à l'introduction de la variable  $X_i$  dans l'équation de régression étant donné que les variables  $X_1 \dots X_{p-1}$  se trouvent déjà dans l'équation de régression.  
 On suppose qu'on renumérote les variables après l'introduction de  $X_i$ .



### 3. 7 Description générale des solutions

Dans le cadre de ce travail on a choisi la méthode de sélection mixte ou de régression pas à pas pour les avantages cités dans 3.5.

Différents critères statistiques peuvent être utilisés pour la sélection des variables qui feront partie de l'équation de régression :

1. Ajouter la variable qui produit la plus grande diminution de la variance résiduelle.

2. Ajouter la variable qui produit la plus grande augmentation de la corrélation multiple entre Y et les variables explicatives sélectionnées.

3. Ajouter la variable explicative pour laquelle son coefficient de corrélation partielle est plus significatif.

4. Eliminer la variable pour laquelle son coefficient de régression est non significatif.

5. Ne pas éliminer une variable explicative si la valeur de  $F$  sortie, est plus grande qu'une valeur spécifiée.

6. Ne pas ajouter une variable si la valeur de  $F$  entrée, est au dessous d'une valeur spécifiée.

La méthode choisie procède à chaque itération de la façon suivante :

- Ajouter la variable qui produit la plus grande diminution de la variance résiduelle d'entre toutes les variables explicatives en considérant le critère 6.
- Eliminer la variable qui produit la plus petite augmentation de la variance résiduelle en considérant le critère 5.
- Elle prévoit aussi la possibilité d'ajouter certaines variables obligatoires à l'équation selon le critère du statisticien sans considérer les règles énoncées précédemment.

### 3. 7.1 Détails de calcul

Le calcul des paramètres statistiques et coefficients de la régression est fait à partir de la matrice des sommes de carrés et des produits des écarts.

L'application de l'opérateur "sweep" à chaque ajout d'une nouvelle variable et l'application de l'opération inverse à chaque élimination, permettent de calculer les coefficients de la régression et la somme des carrés des écarts résiduelles à chaque pas de la régression comme décrit dans le paragraphe 3.4.1.

La somme des carrés des écarts résiduelle pour la variable dépendante  $X_d$  est donnée par  $\hat{a}_{dd}$ .

La réduction de la somme des carrés des écarts résiduelle due à l'introduction de la variable  $X_k$  dans l'équation est :

$$\begin{aligned}
 \text{SCE}_{X_k/X_1 \dots X_{p-1}} &= \text{SCE}_{Y, X_1 \dots X_{p-1}} - \text{SCE}_{Y, X_1 \dots X_k \dots X_p} \\
 &= \hat{a}_{dd} - \left( \hat{a}_{dd} - \frac{\hat{a}_{dk} \hat{a}_{kd}}{\hat{a}_{kk}} \right) \\
 &= \frac{\hat{a}_{kd}^2}{\hat{a}_{kk}}
 \end{aligned}$$

La variable est ajoutée si la diminution de la variance résiduelle est significative. Le test est fait en calculant la variable  $F$  de Snedecor donnée par la formule 3. 37



où

$$SCE_{Y.X_1 \dots X_K \dots X_P} = \hat{a}_{dd} - \frac{\hat{a}_{kd}^2}{\hat{a}_{kk}}$$

$$F_{\text{entrée}} = \frac{(SCE_{X_K / X_1 \dots X_P}) (n-p-2)}{SCE_{Y.X_1 \dots X_K \dots X_{p+1}}} \quad (3.38)$$

Dans le cas d'élimination d'une variable, l'augmentation de la variance résiduelle qui résulte est donnée par :

$$\begin{aligned} SCE_{X_K / X_1 \dots X_P} &= -(SCE_{Y.X_1 \dots X_K \dots X_P} - SCE_{Y.X_1 \dots X_{P-1}}) \\ &= -(\hat{a}_{dd} - (\hat{a}_{dd} - \frac{\hat{a}_{dk} \hat{a}_{kd}}{\hat{a}_{kk}})) \\ &= \frac{\hat{a}_{kd}^2}{\hat{a}_{kk}} \end{aligned}$$

$$F_{\text{sortie}} = \frac{(SCE_{X_K / X_1 \dots X_{p-1}}) (n-p-1)}{SCE_{Y.X_1 \dots X_K \dots X_P}} \quad (3.39)$$

Pour la sortie des résultats on a calculé les valeurs  $T_{\text{entrée}} = \sqrt{F_{\text{entrée}}}$  et  $T_{\text{sortie}} = \sqrt{F_{\text{sortie}}}$ .



Ils ont une distribution T de Student à  $n-p-2$  et  $n-p-1$  degrés de liberté respectivement. Ceci permet à l'utilisateur de trouver le degré de signification des coefficients de régression en en faisant le test sur T plutôt que sur F car la table de T ne comporte qu'un degré de liberté et est souvent plus précise.

### 3. 8 Description de l'algorithme

Les principales étapes pour effectuer la régression multiple sont :

- Lecture et validation des variables qui interviennent dans l'équation de régression.
- Lecture du choix des informations à imprimer.
- Calcul des paramètres statistiques : moyenne, minimum, maximum, SCE, variance, écart-type, SPE, covariance, coefficient de corrélation partielle, pour les variables dépendantes et explicatives.
- Tant qu'il existe des variables obligatoires faire :
  - ajouter une variable à l'équation,
  - calculer et imprimer les paramètres statistiques et les coefficients de l'équation de régression.
- Tant qu'il existe des variables pour entrer dans l'équation, non fortement corrélées avec celles déjà existantes et qui produisent une diminution significative de la variance résiduelle, faire :
  - ajouter une variable à l'équation,
  - calculer et imprimer les paramètres statistiques et les coefficients de la régression,
  - trouver une variable K pour sortir.

Si l'augmentation de la variance résiduelle n'est pas significative, alors :

- éliminer la variable K de l'équation
- calculer et imprimer les paramètres statistiques et les coefficients de l'équation de régression.

Le calcul et l'impression des paramètres statistiques et des coefficients de la régression comprend :

- le calcul des paramètres statistiques qui donnent l'information sur la signification de la régression :

SCERES, ECTRES,  $R^2$ .

- le calcul des coefficients de la régression  $b_i$ , la valeur de T de Student pour tester sa signification, les erreurs standards.
- le calcul des corrélations partielles pour les variables non présentes dans l'équation de régression et T de Student pour tester sa signification.
- Si l'utilisateur à choisi l'impression des paramètres statistiques à chaque pas de la régression alors :
- l'impression des paramètres statistiques.



## Chapitre : 4 : IMPLEMENTATION DU LOGICIEL INTERACTIF POUR EFFECTUER LA REGRESSION MULTIPLE

### 4. 1. Introduction

Le logiciel a été conçu pour donner à l'utilisateur la possibilité de saisir et de préparer les données pour effectuer plusieurs analyses statistiques et en particulier la régression multiple.

Le programme a été écrit en Fortran en considérant que dans le milieu où il pourra être utilisé actuellement on ne possède que ce compilateur et on n'envisage pas de changement à court terme.

La structure du logiciel est modulaire, ce qui donne la possibilité d'ajouter d'autres analyses statistiques qu'utilisent certaines routines de base déjà existantes.

Les opérations sur les données ont été groupées dans un seul module de façon à donner aux traitements une certaine indépendance de la structure de données choisie.

La structure du programme est présentée à la figure (4.2) suivant les notations de JACKSON (1983). La description du programme comporte les spécifications simplifiées de chaque module et la description des routines associées. Les spécifications des paramètres seront reprises dans l'annexe.

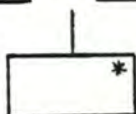
---

(1) Les trois composants élémentaires pour décrire la structure du programme sont :



#### Séquence

Cette structure représente l'ordre d'exécution de traitements, de gauche à droite.



#### Itération

Ce composant représente la répétition d'un traitement.



#### Sélection

Cette structure représente le choix entre deux traitements alternatifs.



Le programme peut résoudre des problèmes de régression avec un maximum de 30 variables et 1000 données par variables. Il est possible de stocker sur un fichier disque 500 variables. Les valeurs limites indiquées ci-dessus peuvent être changées pour les adapter aux possibilités de chaque ordinateur.

Le code source du programme occupe 110 K bytes.

La version exécutable du système REGM occupe en mémoire centrale 30 K mots de 36 bits pour l'ordinateur DEC-20 de l'Institut d'Informatique des Facultés Universitaires N.-D. de la Paix, il peut être adapté à des microordinateurs, en utilisant la technique de "recouvrement" (17).

Cette technique permet de diviser un programme en un module permanent et plusieurs modules rechargeables. Le module permanent est chargé en mémoire centrale pendant toute l'exécution du programme et les modules rechargeables sont chargés quand ils sont nécessaires.

La technique de recouvrement est disponible sur le compilateur "FORTRAN Microsoft version 3.20" qui tourne sur des microordinateurs de 16 bits tels que IBM-PC-XT, KAYPRO 10, avec le système d'exploitation MS-DOS.

Diagramme de Flux du Système REGM

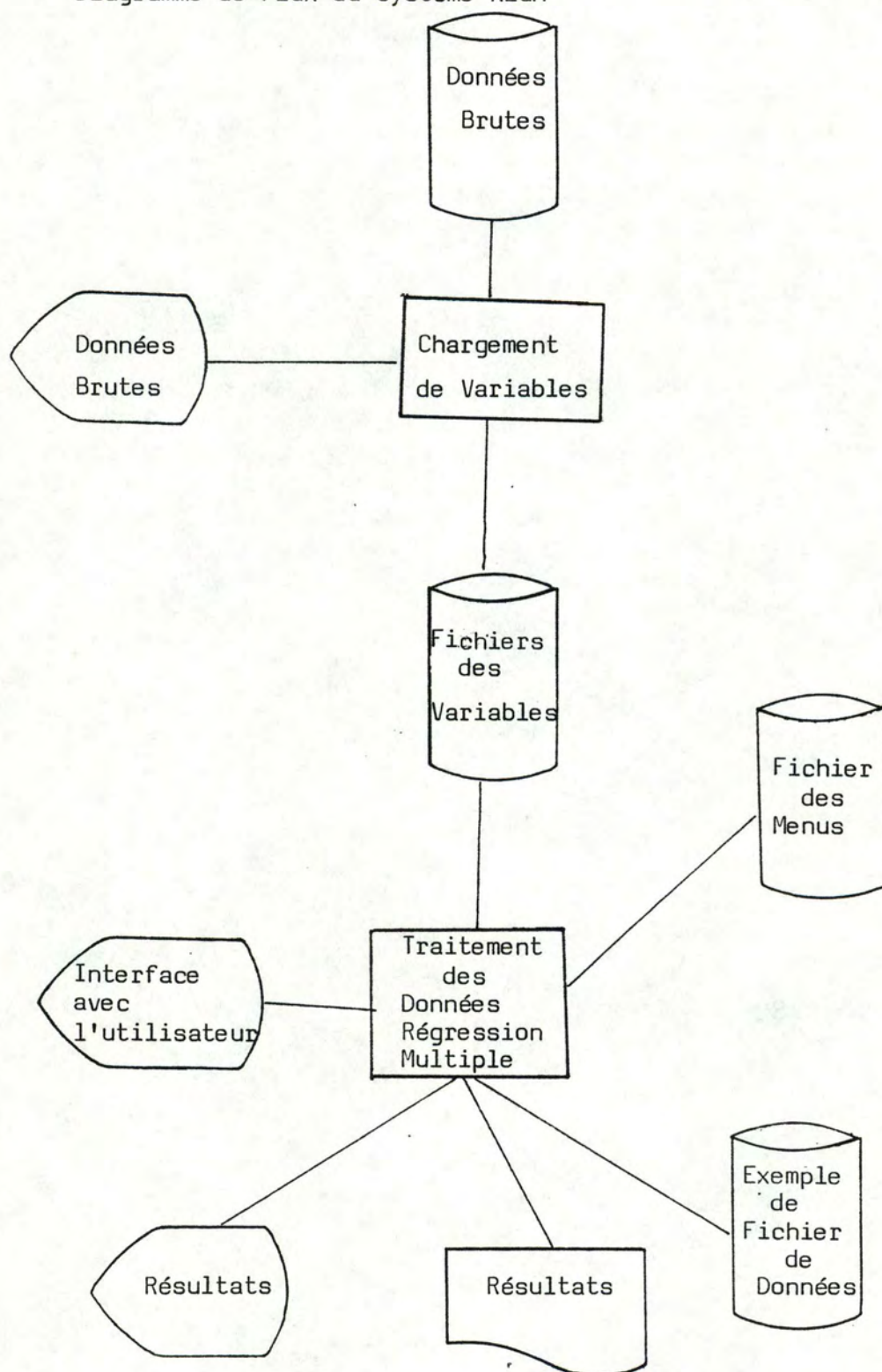
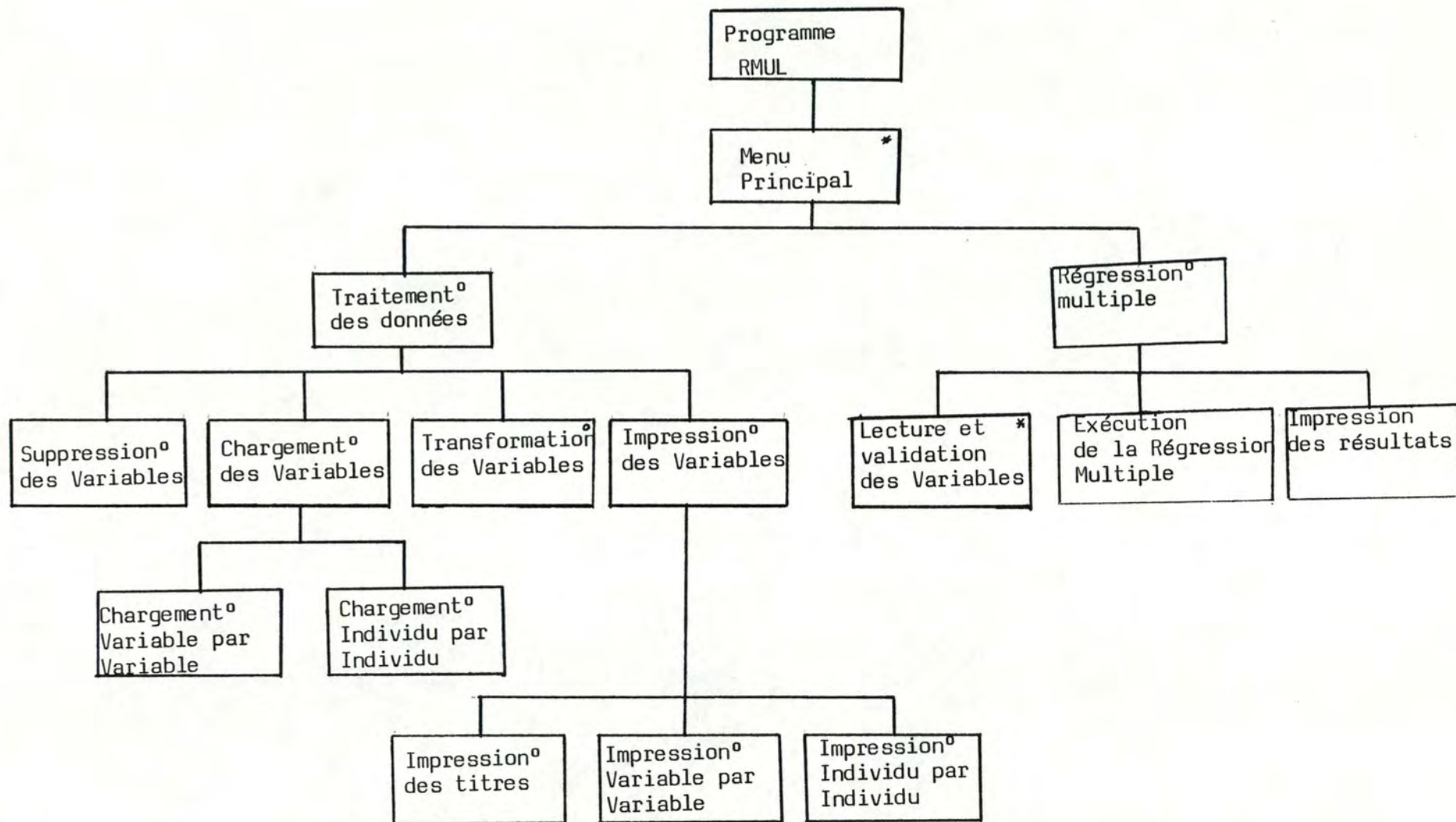


Figure 4.1





STRUCTURE DU PROGRAMME RMUL

Figure 4. 2



#### 4. 2. Module de Gestion de menus

Il permet de gérer deux scénarios pour le choix des options.  
Un scénario avec présentation de menus et un scénario sans présentation de menus.  
Dans chaque étape le module fait la validation du choix de l'utilisateur.

##### Entrée :

- Code qu'indique le menu à afficher (interne)
- Choix de l'utilisateur

##### Sortie :

- Affichage de menu
- Code qu'indique le traitement choisi par l'utilisateur
- Messages d'erreur

##### Opérations sur les données permanentes :

Nom des Fichiers	Description	Opérations
menu.dat	fichier qui contient les menus	Consultation
answer.dat	fichier qui contient les réponses valides dans chaque étape	Consultation

## Structure Physique

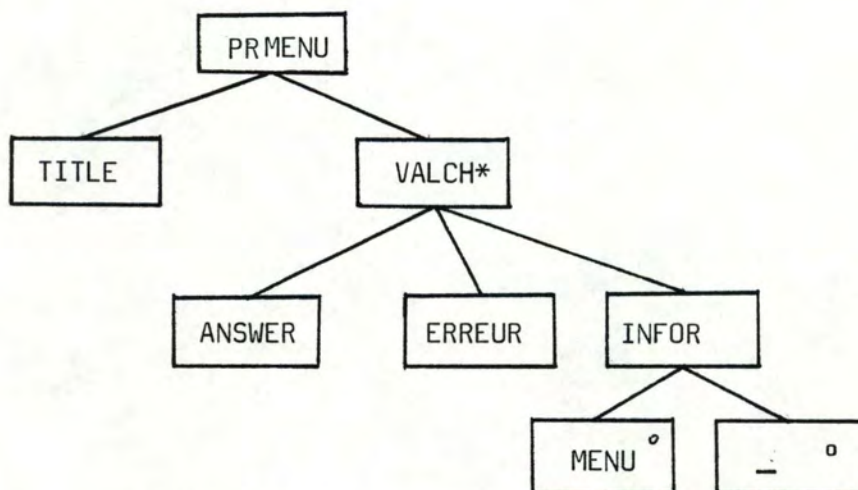


Figure 4.3

## Description

Routine PRMENU

Objectif : - Afficher le titre du menu numéro i (TITLE)  
(i est un argument d'entrée)

Routine VALCH

Objectifs : - Lire et valider le choix de l'utilisateur (ANSWER)

- Afficher un message d'erreur en cas de réponse incorrecte de l'utilisateur (ERREUR)
- Afficher le menu  $i$ , (MENU), si la réponse de l'utilisateur correspond à une demande d'information (INFOR)

#### 4. 3. Modules de traitements de données :

Ce sont les modules qui permettent d'effectuer des opérations sur les données stockées sur des fichiers de variables, tels sont :

- Chargement des variables
- Suppression des variables
- Transformation des variables
- Impression des variables
- Accès aux structures de données

##### 4. 3.1 Modules de chargement des variables

Il permet de charger sur disque les titres, paramètres et données des variables.

Les données peuvent se présenter codées d'une des deux façons suivantes:

- a) Variable par Variable
- b) Individu par Individu

Entrée

- Messages de dialogue interactif
- Pour la saisie de données, il y a deux possibilités :
  - a) à partir d'un fichier créé par l'utilisateur  
à l'aide d'un éditeur de texte
  - b) par terminal lors de l'exécution du module



## Sortie :

- Affichage à l'écran des menus et des messages d'erreur,
- Affichage de (s) numéro (s) assigné(s) par le système aux variables ajoutées.

## Opérations sur les données permanentes :

Nom des Fichiers	Description	Opérations
XX.tit <sup>(2)</sup>	fichier des titres	création-modification
XX.par	fichier des paramètres	création-modification
XX.don	fichier des données	création-modification
XX.aux	fichier qui contient les listes de pointeurs	création-modification-consultation
menu.dat	fichier des menus	consultation
answer.dat	fichier contenant les réponses correctes pour chaque menu	consultations
exreg.dat	exemple du fichier de données pour chargement variable par variable.	consultation
exind.dat	exemple du fichier de données pour chargement individu par individu	consultation

---

(2) XX nom du fichier (maximum 6 caractères) donnés par l'utilisateur

## Structure Physique

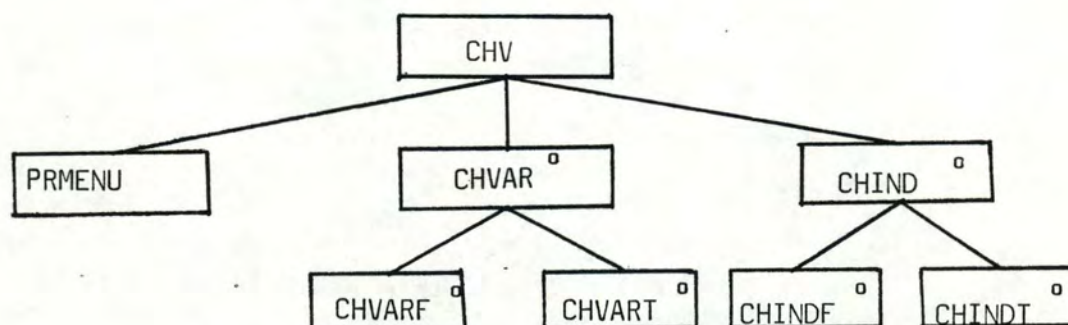


Figure 4.4

## Description

## Routine CHV

Objectif : Présenter le menu du type de chargement et valider le choix de l'utilisateur (PRMENU)

## Routine CHVAR

Objectif : Lire à partir du fichier (CHVARF), ou terminal (CHVART), selon le choix de l'utilisateur, variable par variable leurs titres, paramètres et données et ajouter aux fichiers de variables.

## Routine CHIND

Objectif : Lire à partir du fichier (CHINDF), ou du terminal (CHINDT) les titres de chaque variable, les paramètres et données codées individu par individu et les ajouter au fichier des variables.



#### 4.3.2. Module de Suppression des Variables

Il permet d'éliminer une, plusieurs ou toutes les variables du fichier de variables.

En cas d'élimination de toutes les variables, les fichiers XX.tit, XX.par, XX.don, XX.aux sont effacés.

##### Entrée

- Préfixe du nom des fichiers de variables  
(XX : maximum 6 caractères)
- Nombre et numéros des variables à supprimer

##### Sorties

- Messages d'erreur
- Affichage du titre de la variable à supprimer
- Message de confirmation de l'opération

##### Opérations sur les données permanentes

Nom des fichiers	Opération
XX.tit	Consultation
XX.aux	Consultation, Modification

##### Implémentation

- Ce module est implémenté dans la routine ELIMIN



## Structure Physique

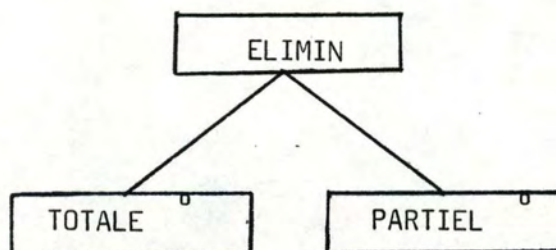


Figure 4.5

## Description

Routine ELIMIN : Cette routine élimine une, plusieurs ou toutes les variables des fichiers de variables.

En cas d'élimination de toute les variables, les fichiers de variables seront effacés.

#### 4.3.3. Module de transformation des variables

Ce module permet d'effectuer des transformations des variables, les résultats étant eux-mêmes transférés sur disque comme de nouvelles variables.

##### Entrée :

- Code du type de transformation à effectuer.
- Numéros des variables qui interviennent dans la transformation.

##### Sortie

- Affichage à l'écran des possibilités de transformation.
- Messages d'erreur.

##### Opérations sur les données permanentes :

Nom des fichiers	Opérations
XX.tit	Consultation, Modification
XX.par	Consultation, Modification
XX.données	Consultation, Modification

##### Implémentation

##### Routine TRVAR

- Objectifs :
- Afficher les possibilités de transformation.
  - Lire et valider les numéros des variables qui interviennent dans la transformation.
  - Lire les données des variables.
  - Effectuer le traitement choisi par l'utilisateur.
  - Ecrire les nouvelles variables sur les fichiers de titres, paramètres et données.

#### 4.3.4. Module d'impression des variables

Le rôle de ce module est d'imprimer sur papier ou à l'écran les titres, paramètres et données d'une ou plusieurs variables enregistrées sur disque.

##### Entrée

- Choix du type d'impression.
- Format d'impression.
- Numéros des variables à imprimer.

##### Sortie

- Messages d'erreur
- Impression du rapport choisi par l'utilisateur, à l'écran ou sur papier.

Il peut être de trois types :

- Impression de titres de toutes les variables.
- Impression variable par variable, des titres, paramètres et données.
- Impression individu par individu, des titres, paramètres et données.

Dans les deux derniers cas on imprime des paramètres statistiques additionnels : moyenne, minimum, maximum et variance.

Opérations sur les données permanentes :

Nom des fichiers	Opérations
XX.tit	Consultation
XX.par	Consultation
XX.don	Consultation



## Structure Physique

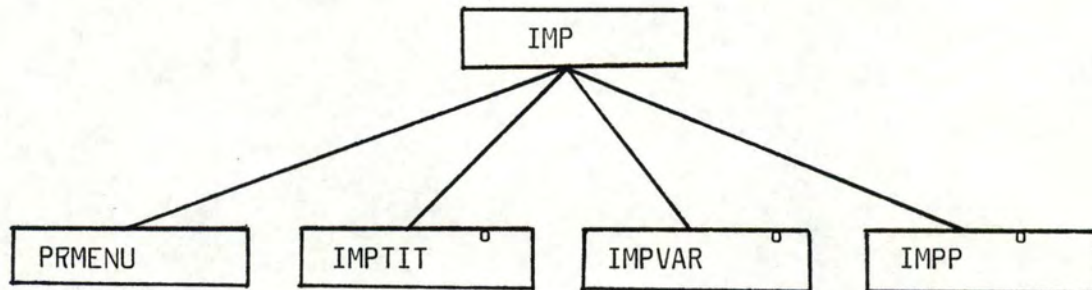


Figure 4.6

## Description

## Routine IMP

- Objectifs :
- Présenter le menu du type d'impression et valider le choix de l'utilisateur (PRMENU)
  - Imprimer les titres de toutes variables (IMPTIT).
  - Imprimer variable par variable, les titres, paramètres et données (IMPVAR).
  - Imprimer les titres, paramètres et données de plusieurs variables simultanément (IMPP).

#### 4.3.5. Module d'accès aux structures de données:

Il permet de manipuler les fichiers des titres, paramètres et données.

Les opérations que l'on peut effectuer sur ces fichiers sont :

1. Ajouter les titres, paramètres et données des variables.
2. Créer les fichiers..
3. Ouvrir les fichiers.
4. Fermer les fichiers.
5. Eliminer les variables.
6. Effacer les fichiers.
7. Consulter les titres, paramètres et données des variables.
8. Modifier les items d'une variable.

Entrée : - Nom du fichier.

- Code indiquant le type de transformation à effectuer.
- Pour les opérations 1, 2, 7 et 8 : numéro de la variable.

Sortie : - Pour les opérations 1, 7, 8 : titres, paramètres et données des variables.

- Code d'erreur = 1 si l'opération demandée s'effectue correctement,

= 0 dans le cas contraire.

Implémentation : Ce module est implémenté dans la routine TYPEAB.



#### 4.4 Module de Régression Multiple

Ce module permet de calculer et imprimer à l'écran ou sur papier les résultats de régressions multiples, progressives et d'écrire sur disque les valeurs estimées d'une variable dépendante calculée à l'aide d'équations de régression multiple.

Entrée : Numéros des variables dépendantes et explicatives qui interviennent dans la régression.

Sortie : Impression des paramètres statistiques à chaque pas de la régression.  
Stockage sur disque de la variable dépendante estimée et des résidus de la régression.

La structure physique est montrée à la figure 4.7.



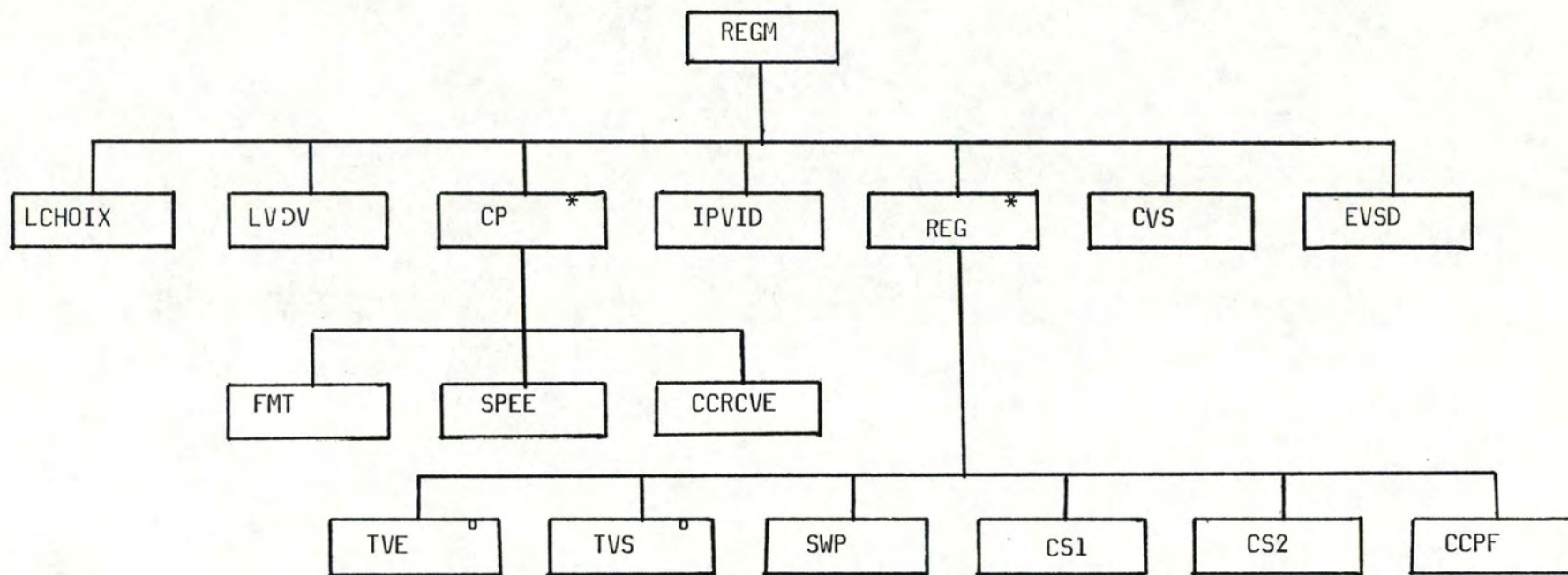


figure 4.7

## Description

## Routine REGM

- Objectifs :
- Lire les choix d'impression des résultats (LCHOIX).
  - Lire et valider les numéros des variables qui interviennent dans la régression multiple (LV DV).
  - Calculer des paramètres statistiques pour les variables dépendantes et indépendantes (CP).
  - Imprimer sur papier ou à l'écran les paramètres statistiques calculés (IPVID).
  - Effectuer la régression multiple (REG).
  - Calculer les valeurs de "Y" estimées et les résidus de la régression (CVS).
  - Enregistrer la variable estimée et les résidus de la régression sur disque (EVSD).

## Routine REG

- Objectifs :
- Choisir une variable à ajouter à l'équation de régression (TVE).
  - Choisir une variable de l'équation de régression (TVS).
  - Calculer les coefficients de l'équation de régression et la somme des carrés des écarts résiduelle (SWP).
  - Calculer les paramètres utiles pour l'analyse de la variance c'est-à-dire : SCERES, CMRES, SCEREG, CMREG, ( cf. 3.2) (CS1).
  - Calculer l'erreur standard  $s_i$  pour les coefficients de la régression et la valeur  $F_i$  de Snedecor pour tester sa signification (CS2).
  - Calculer les coefficients de corrélation partielle et la valeur de F pour tester sa signification (CCPF).

## Routine CP

- Objectifs :
- Lire les données variables dépendantes et explicatives et former la matrice de travail (FMT).
  - Calculer la moyenne, le minimum, maximum et la matrice de sommes des carrés et de produits des écarts (CSPEE).
  - Calculer les coefficients de corrélation, covariance, variance et écart type pour les variables explicatives et les variables dépendantes (CCRCUE).



## Chapitre 5 : MANUEL DE L'UTILISATEUR

### 5. 1. Introduction

Le programme RMUL permet à l'utilisateur d'effectuer une régression multiple en utilisant comme données des variables stockées sur disque.

Différentes options sont disponibles pour manipuler les variables, ainsi :

- saisie des données,
- élimination des variables,
- impression des variables,
- transformation des variables,
- stockage sur disque de la variable estimée et des résidus de la régression.

Pour utiliser ce programme, deux scénarios ont été implémentés pour le choix des options : un scénario avec présentation des menus, et un scénario sans présentation de menus. Dans le deuxième cas on a la possibilité d'obtenir des informations sur les choix possibles en tapant le caractère I.

Dans la suite on utilisera les conventions suivantes :

- = Position du curseur à l'écran
- ( R ) = Pousser sur la touche (Return)

Fichier de variables : fichiers permanents de titres, paramètres et données de variables stockées sur disque.



## 5. 2. Exécution du programme RMUL

Si on dispose d'une version exécutable du programme RMUL on commencera la session en tapant :

RUN RMUL ( R )

Une fois la commande lancée il va apparaître le message suivant :

Bienvenue au programme de régression multiple

Souhaitez-vous des informations sur la manipulation de ce programme (O/N) : ☐

Si la réponse est O (oui) on voit apparaître un résumé des possibilités du programme. Le paragraphe qui suit reproduit le résumé

### 5. 3. Information sur le programme RMUL

Ce programme permet de :

- Créer des fichiers permanents pour stocker les titres, paramètres et données des variables à utiliser pour effectuer des analyses statistiques.
- Charger des variables à partir du terminal ou des fichiers temporaires créés par l'utilisateur.
- Créer des nouvelles variables qui sont les résultats des transformations des variables déjà existantes.
- Supprimer des variables contenues dans le fichier de variables.
- Créer de nouvelles variables qui sont les résultats de transformations de variables déjà existantes.
- Effectuer la régression multiple.
- Imprimer la variable estimée et les résidus sur disque et/ou sur papier.
- Imprimer la ou les variables contenues dans le fichier de variables.

A chaque étape tapez I pour consulter les options possibles.

Si on choisit I à cette étape on a l'affichage du menu principal.



#### 5. 4. Menu principal

Choix d'un type de travail :

C : Chargement des variables

S : Suppression des variables

T : Transformation des variables

W : Impression des variables

R : Régression Multiple

F : Fin

□

L'option F permet de terminer l'exécution du programme

Dans les autres cas on retourne à ce menu principal dès que l'exécution est terminée.

### 5. 5. Chargement des Variables

Le programme demande à l'utilisateur :

- Le nom du fichier de variables (maximum 6 caractères majuscules)

NOM DU FICHIER DE VARIABLES   

Si le fichier n'existe pas le système crée un nouveau fichier et initialise les variables nécessaires pour son exécution.

Ensuite le système affiche le menu de chargement des variables :

TYPE DE CHARGEMENT

V : Variable par Variable

P : Individu par Individu

☐

- L'option V, permet de charger à partir d'un fichier ou du terminal les titres, paramètres et données de variables codées variable par variable.
- L'option P permet de charger à partir d'un fichier ou du terminal les titres, paramètres et données des plusieurs variables de façon simultanée.
- L'utilisateur est informé du numéro xx, assigné à la variable par le message :

LA VARIABLE EST STOCKEE DANS LA POSITION : xx



## 5. 6. Suppression des variables

Le programme demande à l'utilisateur :

- Le nom du fichier de variables (maximum 6 caractères majuscules).

NOM DU FICHIER DE VARIABLES ☐

- Le type d'élimination des variables, total ou partiel :

VOUDRIEZ-VOUS EFFACER TOUTES LES VARIABLES (O/N) : ☐

Si la réponse est O (oui), on efface tout le fichier, et on retourne au menu principal. Dans le cas contraire le programme demande :

- le nombre de variables à supprimer et leur numéro

DONNER LE NOMBRE DE VARIABLES A SUPPRIMER ☐

DONNER LE NUMERO DE LA VARIABLE A SUPPRIMER ☐

- Pour chaque variable le système affiche son titre et le message suivant :

C'EST BIEN CETTE VARIABLE QUE VOUS VOULEZ SUPPRIMER (O/N ) ☐

Si la réponse est N (non), les deux derniers messages sont répétés.

## 5. 7. Impression des variables

Cette option permet d'effectuer l'impression des titres, paramètres et données d'une ou plusieurs variables contenues dans le fichier de variables.

Le programme demande à l'utilisateur : le nom du fichier de variables (maximum 6 caractères majuscules) :

NOM DU FICHIER DE VARIABLES □

Ensuite, on a l'affichage du menu d'impression :

IMPRESSION DES VARIABLES

T : Impression des titres

V : Impression variable par variable

P : Impression individu par individu

F : Fin

□

- L'option T permet l'impression des titres de toutes les variables.
- L'option V permet l'impression variable par variable de titres, paramètres et données.
- L'option P permet l'impression simultanée de titres, paramètres et données de plusieurs variables.
- L'option F termine l'exécution du programme.

Dans chaque cas le système permet de choisir l'impression sur papier ou sur terminale.



## 5. 8. Transformation des variables

Cette option permet à l'utilisateur d'effectuer des transformations de variables et de calculer des fonctions de deux variables à partir de données enregistrées sur disque dans un fichier de variables, les résultats étant eux-mêmes transférés sur disque dans de nouvelles variables.

Le système permet de choisir entre différentes options au moyen du menu :

### POSSIBILITES :

1.  $x_2 = c * x_1$
2.  $x_3 = x_1 + x_2$
3.  $x_3 = x_1 - x_2$
4.  $x_3 = x_1 * x_2$
5.  $x_3 = x_1 / x_2$
6.  $x_2 = x_1 ** c$
7.  $x_3 = F( x_1 )$

### FONCTIONS F DISPONIBLES

ABS = Valeur Absolue

LOGE = Logarithme Népérien

LOGD = Logarithme Décimal

EXP = Exponentielle

NEGA = Suppression des valeurs négatives

NEZE = Suppression des valeurs nulles et négatives

Choix : ☐

### 5. 9. Régression Multiple

Le programme demande à l'utilisateur :

NOMBRE DE VARIABLES OBLIGATOIRES : ☐

NOMBRE DE VARIABLES FACULTATIVES : ☐

Les variables obligatoires seront conservées dans l'équation de régression quelle que soit la signification de leurs coefficients de régression.

Une variable facultative peut être éliminée de l'équation de régression au moment où son coefficient de régression n'est pas significatif, selon le procédé expliqué au chapitre 3. 4. La somme des variables obligatoires et facultatives ne doit pas dépasser "30". Si cette condition est remplie le système demande :

NUMEROS DES VARIABLES OBLIGATOIRES : ☐

puis,

NUMEROS DES VARIABLES FACULTATIVES : ☐

Dans chaque cas l'utilisateur doit taper les numéros qui identifient les variables, séparés par un ou plusieurs espaces.

Si les variables identifiées par ces numéros existent, le système demande :

DANS LES MESSAGES SUIVANTS, CHOISIR ENTRE LES OPTIONS (O/N)

IMPRESSION DES STATISTIQUES A CHAQUE PAS : ☐

IMPRESSION DE LA DERNIERE STATISTIQUE : ☐

STOCKAGE DE LA VARIABLE DEPENDANTE ESTIMEE : ☐

STOCKAGE DES RESIDUS SUR DISQUE : ☐

IMPRESSION DES VALEURS DE Y ESTIMEES ET DE RESIDUS : ☐

VOULEZ-VOUS LES RESULTATS SUR ECRAN OU SUR PAPIER (E/P) : ☐

### 5.10 Exemple d'utilisation

Afin d'illustrer l'utilisation du programme REGM on a pris comme exemple, des données écologiques de la Faculté des Sciences Agronomiques de l'Etat à Gembloux.

Il s'agit d'étudier la relation entre la variable dénommée "Point de Flétrissement" et 10 variables explicatives.

L'exécution du problème a été réalisé deux fois avec des valeurs limites de T de Student pour entrer et sortir les variables de l'équation, égaux à 1 et 2.

Dans les pages qui suivent on présente les données et un exemple d'utilisation du programme depuis le chargement des variables jusqu'à l'impression des résultats de la régression.



\*\*\*\*\*  
\*        EXEMPLE DES DONNEES        \*  
\*\*\*\*\*

# ANALYSE DE DONNEES ECOLOGIQUES

## TITRES DES VARIABLES:

- |    |                                 |
|----|---------------------------------|
| 1  | ARGILE %                        |
| 2  | LIMON %                         |
| 3  | SABLE FIN % (0.05 - 0.1MM)      |
| 4  | SABLE FIN % (0.1 - 0.2 MM)      |
| 5  | SABLE GROSSIER % (0.2 - 0.5 MM) |
| 6  | SABLE GROSSIER % (0.5 - 1 MM)   |
| 7  | SABLE GROSSIER % (1 - 2MM)      |
| 8  | GRAVIERS %                      |
| 9  | POINT FLETRISS. PT %            |
| 10 | MATIERE ORGANIQUE %             |
| 11 | CARBONE ORGANIQUE %             |

## PARAMETRES:

NOMBRE DES DONNEES:        28.

CODE DE FIN DES DONNEES:   30000.

CODE DE DONNEES MANQUANTES: 31000.

## DONNEES:

No :	1	2	3	4	5	6	7	8	9	10	11
1	431.	339.	35.	24.	19.	8.	5.	95.	209.	70.	405.
2	388.	299.	80.	21.	5.	5.	13.	108.	158.	36.	213.
3	364.	380.	27.	16.	11.	8.	7.	18.	325.	24.	724.
4	353.	323.	54.	40.	6.	23.	5.	5.	294.	140.	814.
5	320.	455.	20.	15.	16.	7.	2.	0.	346.	104.	61.
6	272.	237.	56.	56.	64.	62.	88.	272.	157.	88.	513.
7	124.	175.	80.	124.	188.	124.	73.	279.	168.	51.	302.
8	158.	173.	94.	115.	163.	94.	67.	184.	141.	60.	349.
9	184.	262.	94.	84.	127.	38.	65.	92.	202.	81.	471.
10	103.	186.	46.	74.	154.	147.	146.	368.	223.	96.	561.
11	281.	161.	67.	59.	74.	58.	59.	103.	311.	142.	826.
12	287.	247.	67.	60.	68.	43.	29.	88.	234.	103.	604.
13	235.	203.	69.	64.	87.	62.	85.	94.	245.	98.	515.
14	346.	430.	21.	13.	2.	0.	0.	0.	314.	132.	772.
15	291.	334.	83.	60.	27.	15.	4.	0.	250.	101.	575.
16	361.	412.	22.	22.	12.	4.	7.	8.	225.	93.	526.
17	110.	121.	72.	120.	173.	160.	155.	400.	121.	47.	272.
18	117.	50.	117.	231.	237.	99.	74.	372.	62.	16.	95.
19	101.	134.	90.	160.	204.	93.	95.	346.	138.	61.	355.
20	132.	164.	100.	113.	134.	104.	86.	313.	134.	77.	453.
21	237.	152.	110.	159.	86.	75.	11.	2.	237.	80.	469.
22	106.	142.	92.	155.	194.	121.	96.	304.	64.	36.	210.
23	254.	289.	50.	36.	65.	59.	119.	273.	173.	45.	265.
24	290.	389.	51.	40.	30.	15.	2.	117.	244.	94.	548.
25	363.	321.	54.	37.	68.	25.	28.	66.	109.	54.	308.
26	354.	427.	15.	6.	1.	0.	1.	113.	259.	120.	698.
27	139.	167.	63.	109.	139.	113.	94.	244.	249.	131.	765.
28	269.	232.	35.	52.	53.	42.	59.	112.	337.	202.	1180.



@EXEC @PROG.COMD

# BIENVENUE AU PROGRAMME DE REGRESSION MULTIPLE

SOUHAITEZ-VOUS DES INFORMATIONS SUR LA MANIPULATION  
DE CE PROGRAMME? (O/N): 0

## PROGRAMME RMUL =====

CE PROGRAMME PERMET DE:

- CREER DES FICHIERS PERMANENTS POUR STOCKER LES TITRES  
PARAMETRES ET DONNEES DES VARIABLES A UTILISER POUR  
EFFECTUER DES ANALYSES STATISTIQUES.
- CHARGER DES VARIABLES A PARTIR DU TERMINAL OU DES  
FICHIERS TEMPORELS CREEES PAR L'UTILISATEUR.
- CREER DE NOUVELLES VARIABLES QUI SONT LES RESULTATS  
DES TRANSFORMATIONS DES VARIABLES DEJA EXISTANTES.
- EFFECTUER LA REGRESSION MULTIPLE.
- ELIMINER UNE OU PLUSIEURS VARIABLES CONTENUES DANS  
LES FICHIERS PERMANENTS.
- IMPRIMER LE RESULTAT DE LA REGRESSION ET LES RESIDUS  
SUR DISQUE ET/OU SUR PAPIER.
- IMPRIMER LA OU LES VARIABLES CONTENUES DANS LES  
FICHIERS DES DONNEES PERMANENTS.

L'EXECUTION DE LA REGRESSION MULTIPLE REQUIERT  
L'EXISTENCE D'UN FICHIER DES DONNEES PERMANENT CREE  
PAR CE PROGRAMME LORS DU CHARGEMENT DES VARIABLES.

DANS CHAQUE ETAPE TAPPEZ I SI VOUS SOUHAITEZ  
CONSULTER LES OPTIONS POSSIBLES

### CHOIX D'UN TYPE DE TRAVAIL

# I

C: CHARGEMENT DES VARIABLES  
S: SUPPRESSION DES VARIABLES  
T: TRANSFORMATION DES VARIABLES  
W: IMPRESSION DES VARIABLES  
R: REGRESSION MULTIPLE  
F: FIN

# C



NOM DU FICHIER DES VARIABLES: REGM

CREATION D'UN NOUVEAU FICHIER

TYPE DE CHARGEMENT

# I

V: VARIABLE PAR VARIABLE

P: INDIVIDU PAR INDIVIDU

# V

MANIERE D'INTRODUIRE LES DONNEES

F: PAR FICHIER

E: PAR ECRAN

# F

SOUHAITEZ-VOUS DES INFORMATIONS SUR LE FICHIER  
A CREER PAR L'UTILISATEUR ? (O/N):N

DONNEZ LE NOM DU FICHIER DES DONNEES: DON.DAT  
NOMBRE DES VARIABLES A CHARGER:11

CHARGEMENT DES VARIABLES

TITRE:

ANALYSE DE DONNEES ECOLOGIQUES

ARGILE %

LA VARIABLE EST STOCKEE DANS LA POSITION 1

TITRE:

ANALYSE DE DONNEES ECOLOGIQUES

LIMON %

LA VARIABLE EST STOCKEE DANS LA POSITION 2

TITRE:

ANALYSE DE DONNEES ECOLOGIQUES

SABLE FIN % (0.05 - 0.1MM)

LA VARIABLE EST STOCKEE DANS LA POSITION 3

TITRE:

ANALYSE DE DONNEES ECOLOGIQUES

SABLE FIN % (0.1 - 0.2 MM)

LA VARIABLE EST STOCKEE DANS LA POSITION 4

TITRE:

ANALYSE DE DONNEES ECOLOGIQUES

SABLE GROSSIER % (0.2 - 0.5 MM)

LA VARIABLE EST STOCKEE DANS LA POSITION 5

TITRE:

ANALYSE DE DONNEES ECOLOGIQUES

SABLE GROSSIER % (0.5 - 1 MM)

LA VARIABLE EST STOCKEE DANS LA POSITION 6

TITRE:

ANALYSE DE DONNEES ECOLOGIQUES

SABLE GROSSIER % (1 - 2MM)

LA VARIABLE EST STOCKEE DANS LA POSITION 7

TITRE:

ANALYSE DE DONNEES ECOLOGIQUES

GRAVIERS %

LA VARIABLE EST STOCKEE DANS LA POSITION 8

TITRE:

ANALYSE DE DONNEES ECOLOGIQUES

POINT FLETRISS. PT %

LA VARIABLE EST STOCKEE DANS LA POSITION 9  
NOM DU FICHER DES VARIABLES : REGM



TITRE:

ANALYSE DE DONNEES ECOLOGIQUES

MATIERE ORGANIQUE %

LA VARIABLE EST STOCKEE DANS LA POSITION 10

TITRE:

ANALYSE DE DONNEES ECOLOGIQUES

CARBONE ORGANIQUE %

LA VARIABLE EST STOCKEE DANS LA POSITION 11

VOUDRIEZ-VOUS CONTINUER L'EXECUTION  
DU PROGRAMME (O/N): 0

CHOIX D'UN TYPE DE TRAVAIL

\* W

NOM DU FICHIER DES VARIABLES : REGM

IMPRESSION DES VARIABLES

\* I

T: IMPRESSION DES TITRES

V: IMPRESSION VARIABLE PAR VARIABLE

P: IMPRESSION INDIVIDU PAR INDIVIDU

F: FIN

\* P

VOUDRIEZ-VOUS L'IMPRESSION A L'ECRAN  
OU SUR PAPIER (P/E): E

NOMBRE DE VARIABLES A IMPRIMER: 6

NUMEROS DES VARIABLES:

1 2 3 4 5 6

VOUDRIEZ-VOUS INDIQUER UN FORMAT D'IMPRESSION  
POUR LES DONNEES (O/N): N

\*\*\*\*\*  
 \* IMPRESSION SIMULTANEE DE PLUSIEURS VARIABLES \*  
 \*\*\*\*\*

# ANALYSE DE DONNEES ECOLOGIQUES

- 1 ARGILE %
- 2 LIMON %
- 3 SABLE FIN % (0.05 - 0.1MM)
- 4 SABLE FIN % (0.1 - 0.2 MM)
- 5 SABLE GROSSIER % (0.2 - 0.5 MM)
- 6 SABLE GROSSIER % (0.5 - 1 MM)

## PARAMETRES:

NOMBRE DES DONNEES: 28.

CODE DE FIN DES DONNEES: 30000.

CODE DE DONNEES MANQUANTES: 31000.

VOUDRIEZ-VOUS INDIQUER UN FORMAT D'IMPRESSION  
 POUR LES DONNEES (O/N): N

NUM	1	2	3	4	5	6
1	431.00	339.00	35.00	24.00	19.00	8.00
2	388.00	299.00	80.00	21.00	5.00	5.00
3	364.00	380.00	27.00	16.00	11.00	8.00
4	353.00	323.00	54.00	40.00	6.00	23.00
5	320.00	455.00	20.00	15.00	16.00	7.00
6	272.00	237.00	56.00	56.00	64.00	62.00
7	124.00	175.00	80.00	124.00	188.00	124.00
8	158.00	173.00	94.00	115.00	163.00	94.00
9	184.00	262.00	94.00	84.00	127.00	38.00



NUM	1	2	3	4	5	6
10	103.00	186.00	46.00	74.00	154.00	147.00
11	281.00	161.00	67.00	59.00	74.00	58.00
12	287.00	247.00	67.00	60.00	68.00	43.00
13	235.00	203.00	69.00	64.00	87.00	62.00
14	346.00	430.00	21.00	13.00	2.00	0.00
15	291.00	334.00	83.00	60.00	27.00	15.00
16	361.00	412.00	22.00	22.00	12.00	4.00
17	110.00	121.00	72.00	120.00	173.00	160.00
18	117.00	50.00	117.00	231.00	237.00	99.00
19	101.00	134.00	90.00	160.00	204.00	93.00
20	132.00	164.00	100.00	113.00	134.00	104.00
21	237.00	152.00	110.00	159.00	86.00	75.00
22	106.00	142.00	92.00	155.00	194.00	121.00
23	254.00	289.00	50.00	36.00	65.00	59.00
24	290.00	389.00	51.00	40.00	30.00	15.00
25	363.00	321.00	54.00	37.00	68.00	25.00
26	354.00	427.00	15.00	6.00	1.00	0.00
27	139.00	167.00	63.00	109.00	139.00	113.00
28	269.00	232.00	35.00	52.00	53.00	42.00
NUM	1	2	3	4	5	6
EFF	28	28	28	28	28	28
MIN	101.00	50.00	15.00	6.00	1.00	0.00
MAX	431.00	455.00	117.00	231.00	237.00	160.00
MOY	248.93	257.29	63.00	73.75	85.96	57.29
ECT	103.47	110.23	28.80	56.05	72.23	48.40
CV	0.42	0.43	0.46	0.76	0.84	0.84

VOUDRIEZ-VOUS CONTINUER L'EXECUTION  
DU PROGRAMME (O/N): 0

# CHOIX D'UN TYPE DE TRAVAIL

# R

NOM DU FICHIER DES VARIABLES : REGM

## REGRESSION MULTIPLE =====

VALEUR LIMITE DE T POUR ENTRER UNE VARIABLE  
DANS L'EQUATION: 1

VALEUR LIMITE DE T POUR SORTIR UNE VARIABLE  
DE L'EQUATION DE REGRESSION: 1

NOMBRE DES VARIABLES OBLIGATOIRES: 0

NOMBRE DES VARIABLES FACULTATIVES: 10

NUMEROS DES VARIABLES FACULTATIVES:

1 2 3 4 5 6 7 8 10 11

NOMBRE DES VARIABLES DEPENDANTES: 1

NUMEROS DES VARIABLES DEPENDANTES: 9

CHOISIR ENTRE LES OPTIONS SUIVANTES: (O/N)

IMPRESSION DES STATISTIQUES A CHAQUE PAS: 0

IMPRESSION DE LA DERNIERE STATISTIQUE: N

STOCKAGE DE LA VARIABLE DEPENDANTE ESTIMEE: 0

STOCKAGE DES RESIDUS SUR DISQUE: 0

IMPRESSION DES VALEURS DE Y EST. ET RESIDUS: 0

VOUDRIEZ-VOUS L'IMPRESSION DES RESULTATS

A L'ECRAN OU SUR PAPIER (E/P): E

VOUDRIEZ VOUS L'IMPRESSION DES PARAMETRES

POUR LES VARIABLES INDEPENDANTES (O/N): 0



# VARIABLES EXPLICATIVES

- 1 ARGILE %
- 2 LIMON %
- 3 SABLE FIN % (0.05 - 0.1MM)
- 4 SABLE FIN % (0.1 - 0.2 MM)
- 5 SABLE GROSSIER % (0.2 - 0.5 MM)
- 6 SABLE GROSSIER % (0.5 - 1 MM)
- 7 SABLE GROSSIER % (1 - 2MM)
- 8 GRAVIERS %
- 10 MATIERE ORGANIQUE %
- 11 CARBONE ORGANIQUE %

NO	MOYENNE	MINIMUM	MAXIMUM	SCE	VARIANCE	ECART-TYPE
1	0.25E+03	0.10E+03	0.43E+03	0.28904187E+06	0.10705254E+05	103.47
2	0.26E+03	0.50E+02	0.46E+03	0.32807772E+06	0.12151027E+05	110.23
3	0.63E+02	0.15E+02	0.12E+03	0.22388000E+05	0.82918520E+03	28.80
4	0.74E+02	0.60E+01	0.23E+03	0.84825250E+05	0.31416759E+04	56.05
5	0.86E+02	0.10E+01	0.24E+03	0.14086496E+06	0.52172209E+04	72.23
6	0.57E+02	0.00E+00	0.16E+03	0.63247714E+05	0.23425079E+04	48.40
7	0.53E+02	0.00E+00	0.16E+03	0.59812108E+05	0.22152632E+04	47.07
8	0.16E+03	0.00E+00	0.40E+03	0.47986571E+06	0.17772804E+05	133.31
10	0.85E+02	0.16E+02	0.20E+03	0.46873857E+05	0.17360688E+04	41.67
11	0.49E+03	0.61E+02	0.12E+04	0.16869367E+07	0.62479137E+05	249.96



NO	NO	SPE	COV	R
1	2	0.25111958E+06	0.93007251E+04	0.8155
1	3	-0.51730001E+05	-0.19159260E+04	-0.6431
1	4	-0.12935250E+06	-0.47908334E+04	-0.8261
1	5	-0.18863707E+06	-0.69865582E+04	-0.9349
1	6	-0.12405843E+06	-0.45947567E+04	-0.9175
1	7	-0.10617864E+06	-0.39325423E+04	-0.8075
1	8	-0.29872443E+06	-0.11063868E+05	-0.8021
1	10	0.27749144E+05	0.10277461E+04	0.2384
1	11	0.18654022E+06	0.69088970E+04	0.2671
2	3	-0.67913001E+05	-0.25152963E+04	-0.7924
2	4	-0.14552100E+06	-0.53896667E+04	-0.8723
2	5	-0.18764071E+06	-0.69496561E+04	-0.8728
2	6	-0.12373229E+06	-0.45826773E+04	-0.8590
2	7	-0.10262143E+06	-0.38007937E+04	-0.7326
2	8	-0.28581329E+06	-0.10585677E+05	-0.7203
2	10	0.30992428E+05	0.11478677E+04	0.2499
2	11	0.14244415E+06	0.52757091E+04	0.1915
3	4	0.36341000E+05	0.13459630E+04	0.8339
3	5	0.39729000E+05	0.14714444E+04	0.7075
3	6	0.21075000E+05	0.78055556E+03	0.5601
3	7	0.13504000E+05	0.50014816E+03	0.3690
3	8	0.44036000E+05	0.16309630E+04	0.4249
3	10	-0.13751000E+05	-0.50929631E+03	-0.4245
3	11	-0.77266001E+05	-0.28617037E+04	-0.3976
4	5	0.98331750E+05	0.36419167E+04	0.8996
4	6	0.55372000E+05	0.20508148E+04	0.7560
4	7	0.36918750E+05	0.13673611E+04	0.5183
4	8	0.12795000E+06	0.47388889E+04	0.6342
4	10	-0.24002500E+05	-0.88898148E+03	-0.3807
4	11	-0.13919675E+06	-0.51554352E+04	-0.3680
5	6	0.83650287E+05	0.30981588E+04	0.8862
5	7	0.69162679E+05	0.25615807E+04	0.7535
5	8	0.21122729E+06	0.78232329E+04	0.8124
5	10	-0.33271929E+05	-0.12322937E+04	-0.4095
5	11	-0.19689840E+06	-0.72925332E+04	-0.4039
6	7	0.53755571E+05	0.19909471E+04	0.8740
6	8	0.14712271E+06	0.54489894E+04	0.8445
6	10	-0.14493572E+05	-0.53679895E+03	-0.2662
6	11	-0.83942859E+05	-0.31089948E+04	-0.2570
7	8	0.14727057E+06	0.54544656E+04	0.8693
7	10	-0.11439358E+05	-0.42367991E+03	-0.2160
7	11	-0.65715536E+05	-0.24339088E+04	-0.2069
8	10	-0.58913573E+05	-0.21819842E+04	-0.3928
8	11	-0.32916186E+06	-0.12191180E+05	-0.3658
10	11	0.22623179E+06	0.83789551E+04	0.8045



VOUDRIEZ VOUS L'IMPRESSION DES PARAMETRES POUR  
LA VARIABLE DEPENDANTE(O/N): 0

VARIABLES DEPENDANTES

9 POINT FLETRISS. PT %

NO	MOYENNE	MINIMUM	MAXIMUM	SCE	VARIANCE	ECART--TYPE
9	0.21E+03	0.62E+02	0.35E+03	0.17121325E+06	0.63412316E+04	79.63

NO	NO	SPE	COV	R
9	1	0.11534351E+06	0.42719817E+04	0.5185
9	2	0.13634901E+06	0.50499632E+04	0.5753
9	3	-0.39464001E+05	-0.14616297E+04	-0.6374
9	4	-0.75462752E+05	-0.27949167E+04	-0.6262
9	5	-0.10587525E+06	-0.39213056E+04	-0.6817
9	6	-0.56514001E+05	-0.20931112E+04	-0.5431
9	7	-0.50642252E+05	-0.18756390E+04	-0.5004
9	8	-0.19693401E+06	-0.72938521E+04	-0.6871
9	10	0.64443500E+05	0.23867963E+04	0.7194
9	11	0.36627425E+06	0.13565713E+05	0.6815

PHASE 1 VARIABLE ENTRANTE = 10 T= 5.28

DLRES 26. SCERES= 0.82614513E+05 ECTRES= 56.37 R\*\*2= 0.4989

NO	B	ECTB	T
10	1.374828	0.260361	5.28
0	94.791403		
1			3.00
2			3.64
3			3.11
4			3.28
5			3.86
6			3.09
7			2.95
8			4.09
11			1.29

PHASE 2 VARIABLE ENTRANTE = 8 T= 4.09

DLRES 25. SCERES= 0.49492409E+05 ECTRES= 44.49 R\*\*2= 0.6878

NO	B	ECTB	T
8	-0.285688	0.069845	4.09
10	1.015760	0.223474	4.55
0	169.986824		
1			0.13
2			1.34
3			2.47
4			1.51
5			1.24
6			0.04
7			0.58
11			1.26



PHASE 3 VARIABLE ENTRANTE = 3 T= 2.47

DLRES 24. SCERES= 0.39437407E+05 ECTRES= 40.54 R\*\*2= 0.7409

NO	R	ECTB	T
3	--0.778509	0.314718	2.47
8	--0.234366	0.066930	3.50
10	0.851879	0.214107	3.98
0	224.953560		
1			1.99
2			1.12
4			0.60
5			0.44
6			1.45
7			0.88
11			1.18

PHASE 4 VARIABLE ENTRANTE = 1 T= 1.99

DLRES 23. SCERES= 0.33656608E+05 ECTRES= 38.25 R\*\*2= 0.7692

NO	R	ECTB	T
1	--0.311153	0.156550	1.99
3	--1.269813	0.386401	3.29
8	--0.404277	0.106288	3.80
10	0.678398	0.220094	3.08
0	374.673687		
2			0.99
4			0.34
5			1.58
6			0.28
7			0.18
11			1.38

PHASE 5 VARIABLE ENTRANTE = 5 T= 1.58

DLRES 22. SCERES= 0.30220798E+05 ECTRES= 37.06 R\*\*2= 0.7834

NO	R	ECTB	T
1	-0.583278	0.229375	2.54
3	-1.067015	0.395728	2.70
5	-0.558130	0.352909	1.58
8	-0.361580	0.106461	3.40
10	0.556481	0.226751	2.45
0	481.314857		
2			2.37
4			0.66
6			0.47
7			0.01
11			1.36

PHASE 6 VARIABLE ENTRANTE = 2 T= 2.37

DLRES 21. SCERES= 0.23833969E+05 ECTRES= 33.69 R\*\*2= 0.8210

NO	B	ECTB	T
1	-0.771283	0.223048	3.46
2	-0.426883	0.179951	2.37
3	-1.937056	0.513713	3.77
5	-1.018405	0.374896	2.72
8	-0.493988	0.111713	4.42
10	0.201663	0.254662	0.79
0	783.203697		
4			0.38
6			1.68
7			1.65
11			0.53



PHASE 7 VARIABLE SORTANTE = 10 T= 0.79

DLRES 22. SCERES= 0.24545673E+05 ECTRES= 33.40 R\*\*2= 0.8241

NO	R	ECTB	T
1	-0.875176	0.178847	4.89
2	-0.510578	0.144402	3.54
3	-2.179056	0.409413	5.32
5	-1.178542	0.312992	3.77
8	-0.540575	0.094158	5.74
0	884.048050		
4			0.40
6			1.92
7			1.85
10			0.79
11			0.87

PHASE 8 VARIABLE ENTRANTE = 6 T= 1.92

DLRES 21. SCERES= 0.20882450E+05 ECTRES= 31.53 R\*\*2= 0.8432

NO	R	ECTB	T
1	-1.125982	0.213504	5.27
2	-0.715156	0.173049	4.13
3	-2.545436	0.431083	5.90
5	-1.244077	0.297453	4.18
6	-0.867620	0.452042	1.92
8	-0.490081	0.092704	5.29
0	1069.642227		
4			0.13
7			1.58
10			0.00
11			0.10

PHASE 9 VARIABLE ENTRANTE = 7 T= 1.58

DLRES 20. SCERES= 0.18562429E+05 ECTRES= 30.47 R\*\*2= 0.8536

NO	B	ECTB	T
1	-1.232113	0.216914	5.68
2	-0.807308	0.177051	4.56
3	-2.787751	0.443774	6.28
5	-1.391218	0.302063	4.61
6	-0.734055	0.444813	1.65
7	-0.532521	0.336817	1.58
8	-0.401550	0.105625	3.80
0	1154.250107		
4			1.20
10			0.38
11			0.32

PHASE 10 VARIABLE ENTRANTE = 4 T= 1.20

DLRES 19. SCERES= 0.17261078E+05 ECTRES= 30.14 R\*\*2= 0.8767

NO	B	ECTB	T
1	-1.325943	0.228477	5.80
2	-0.934113	0.204716	4.56
3	-2.698045	0.445405	6.06
4	-0.515408	0.430637	1.20
5	-1.213735	0.333618	3.64
6	-0.753753	0.440388	1.71
7	-0.878792	0.441304	1.99
8	-0.372107	0.107358	3.47
0	1242.102829		
10			0.76
11			0.80

IMPRESSION DE LA VARIABLE ESTIMEE SUR DISQUE  
 DONNEZ LA QUATRIEME LIGNE DU TITRE POUR LA VARIABLE ESTIMEE  
 VARIABLE ESTIMEE

CHARGEMENT DES VARIABLES  
 TITRE:  
 ANALYSE DE DONNEES ECOLOGIQUES

POINT FLETRISS. PT %  
 VARIABLE ESTIMEE

LA VARIABLE EST STOCKEE DANS LA POSITION 12



IMPRESSION DES VALEURS OBSERVEES ET CALCULEES  
DE LA VARIABLE DEPENDANTE

NUMERO	Y	Y EST	PESIDUS
1	0.209E+03	0.178E+03	0.307E+02
2	0.158E+03	0.160E+03	-0.222E+01
3	0.325E+03	0.291E+03	0.338E+02
4	0.294E+03	0.275E+03	0.189E+02
5	0.346E+03	0.305E+03	0.414E+02
6	0.157E+03	0.177E+03	-0.201E+02
7	0.168E+03	0.145E+03	0.232E+02
8	0.141E+03	0.162E+03	-0.211E+02
9	0.202E+03	0.182E+03	0.197E+02
10	0.223E+03	0.207E+03	0.164E+02
11	0.311E+03	0.284E+03	0.268E+02
12	0.234E+03	0.246E+03	-0.120E+02
13	0.245E+03	0.260E+03	-0.147E+02
14	0.314E+03	0.316E+03	-0.187E+01
15	0.250E+03	0.242E+03	0.819E+01
16	0.225E+03	0.281E+03	-0.562E+02
17	0.121E+03	0.111E+03	0.952E+01
18	0.620E+02	0.398E+02	0.222E+02
19	0.138E+03	0.128E+03	0.102E+02
20	0.134E+03	0.153E+03	-0.188E+02
21	0.237E+03	0.236E+03	0.119E+01
22	0.640E+02	0.117E+03	-0.526E+02
23	0.173E+03	0.152E+03	0.206E+02
24	0.244E+03	0.243E+03	0.102E+01
25	0.109E+03	0.146E+03	-0.366E+02
26	0.259E+03	0.286E+03	-0.271E+02
27	0.249E+03	0.248E+03	0.640E+00
28	0.337E+03	0.358E+03	-0.210E+02

VOUDRIEZ-VOUS CONTINUER L'EXECUTION  
DU PROGRAMME (O/N): 0

CHOIX D'UN TYPE DE TRAVAIL

\* R

NOM DU FICHIER DES VARIABLES  
REGM

REGRESSION MULTIPLE  
=====

VALEUR LIMITE DE T POUR ENTRER UNE VARIABLE  
DANS L'EQUATION: 2

VALEUR LIMITE DE T POUR SORTIR UNE VARIABLE  
DE L'EQUATION DE REGRESSION: 2

NOMBRE DES VARIABLES OBLIGATOIRES: 0

NOMBRE DES VARIABLES FACULTATIVES: 10

NUMEROS DES VARIABLES FACULTATIVES:

1 2 3 4 5 6 7 8 10 11

NOMBRE DES VARIABLES DEPENDANTES: 1

NUMEROS DES VARIABLES DEPENDANTES:

9

CHOISIR ENTRE LES OPTIONS SUIVANTES: (O/N)

IMPRESSION DES STATISTIQUES A CHAQUE PAS: 0

IMPRESSION DE LA DERNIERE STATISTIQUE: N

STOCKAGE DE LA VARIABLE DEPENDANTE ESTIMEE: N

STOCKAGE DES RESIDUS SUR DISQUE: N

IMPRESSION DES VALEURS DE Y EST. ET RESIDUS: 0

VOUDRIEZ-VOUS L'IMPRESSION DES RESULTATS

A L'ECRAN OU SUR PAPIER (E/P): E

VOUDRIEZ VOUS L'IMPRESSION DES PARAMETRES

POUR LES VARIABLES INDEPENDANTES (O/N): N

VOUDRIEZ VOUS L'IMPRESSION DES PARAMETRES POUR

LA VARIABLE DEPENDANTE(O/N): N



PHASE 1 VARIABLE ENTRANTE = 10 T= 5.28

DLRES 26. SCERES= 0.82614513E+05 ECTRES= 56.37 R\*\*2= 0.4989

NO	B	ECTB	T
10	1.374828	0.260361	5.28
0	94.791403		
1			3.00
2			3.64
3			3.11
4			3.28
5			3.86
6			3.09
7			2.95
8			4.09
11			1.29

PHASE 2 VARIABLE ENTRANTE = 8 T= 4.09

DLRES 25. SCERES= 0.49492409E+05 ECTRES= 44.49 R\*\*2= 0.6878

NO	B	ECTB	T
8	-0.285688	0.069845	4.09
10	1.015760	0.223474	4.55
0	169.986824		
1			0.13
2			1.34
3			2.47
4			1.51
5			1.24
6			0.04
7			0.58
11			1.26

PHASE 3 VARIABLE ENTRANTE = 3 T= 2.47

DLRES 24. SCERES= 0.39437407E+05 ECTRES= 40.54 R\*\*2= 0.7409

NO	B	ECTB	T
3	-0.778509	0.314718	2.47
8	-0.234366	0.066930	3.50
10	0.851879	0.214107	3.98
0	224.953560		
1			1.99
2			1.12
4			0.60
5			0.44
6			1.45
7			0.88
11			1.18



IMPRESSION DES VALEURS OBSERVEES ET CALCULEES  
DE LA VARIABLE DEPENDANTE

NUMERO	Y	Y EST	RESIDUS
1	0.209E+03	0.184E+03	0.253E+02
2	0.158E+03	0.153E+03	0.478E+01
3	0.325E+03	0.290E+03	0.351E+02
4	0.294E+03	0.283E+03	0.112E+02
5	0.346E+03	0.309E+03	0.368E+02
6	0.157E+03	0.181E+03	-0.235E+02
7	0.166E+03	0.139E+03	0.285E+02
8	0.141E+03	0.161E+03	-0.200E+02
9	0.202E+03	0.185E+03	0.170E+02
10	0.273E+03	0.218E+03	0.473E+01
11	0.311E+03	0.267E+03	0.440E+02
12	0.234E+03	0.233E+03	0.949E+00
13	0.245E+03	0.271E+03	-0.260E+02
14	0.314E+03	0.314E+03	0.429E+00
15	0.250E+03	0.246E+03	0.384E+01
16	0.225E+03	0.291E+03	-0.663E+02
17	0.121E+03	0.149E+03	-0.280E+02
18	0.620E+02	0.208E+02	0.412E+02
19	0.138E+03	0.104E+03	0.343E+02
20	0.134E+03	0.140E+03	-0.576E+01
21	0.237E+03	0.257E+03	-0.199E+02
22	0.640E+02	0.125E+03	-0.613E+02
23	0.173E+03	0.181E+03	-0.806E+01
24	0.244E+03	0.222E+03	0.221E+02
25	0.109E+03	0.169E+03	-0.600E+02
26	0.259E+03	0.261E+03	-0.227E+01
27	0.249E+03	0.244E+03	0.487E+01
28	0.337E+03	0.331E+03	0.610E+01

VOUDRIEZ-VOUS CONTINUER L'EXECUTION  
DU PROGRAMME (O/N): N



### 5.11 Analyse de résultats de la régression

Dans le premier cas on a utilisé des valeurs de  $T_{\text{entrée}}$  et  $T_{\text{sortie}}$  égaux à 1, ceci permet d'introduire dans l'équation des variables non nuisibles, c'est-à-dire des variables qui n'augmentent pas la variance résiduelle, mais pour lesquelles leurs coefficients de régression ne sont pas nécessairement significatifs à une phase de régression mais qui peuvent devenir significatifs après l'ajoute de nouvelles variables.

Exemple dans la phase n° 4 le coefficient de régression de la variable n° 1 n'est pas significatif au niveau  $\alpha = 0.05$ , mais il devient significatif à la phase n° 5 lors de l'ajoute d'une nouvelle variable.

D'autre part, à la phase n° 6 lors de l'ajoute de la variable n° 2, le coefficient de régression de la variable n° 10 qui était significatif au niveau  $\alpha = 0.05$ , devient non significatif. Pour cette raison cette variable est éliminée de l'équation de régression à la phase n° 7.

On voit également qu'à la phase n° 7 tous les coefficients de l'équation de régression deviennent hautement significatifs ( $\alpha = 0.01$ ).

Le résultat de la phase n° 7 n'a pas été trouvé dans le deuxième essai où on a augmenté les valeurs de  $T_{\text{entrée}}$  et  $T_{\text{sortie}}$  en leur donnant la valeur 2.

L'exemple choisit permet de visualiser la mécanique de la méthode.

En pratique on n'introduit pas simultanément dans une même équation plus de trois ou quatre variables explicatives.

La sélection de l'équation résultante dépendra du critère du statisticien en considérant différents facteurs tels que la difficulté d'obtenir l'information des variables.

## CONCLUSIONS

Le but qui était d'implémenter une version interactive pour faire l'analyse de la régression multiple a été atteint.

Le programme a été testé en utilisant des données de problèmes dont la solution est connue.

Le logiciel facile à utiliser, permet le chargement de données variable par variable et individu par individu.

Les données peuvent se présenter comme des entiers, des réels ou en notation exponentielle selon les formats acceptés par le compilateur FORTRAN.

L'édition des résultats est claire et présente une variété de paramètres utiles pour le statisticien.

Quant aux difficultés rencontrées, elles étaient liées au fait qu'il fallait étudier le système existant qui travaille en mode batch pour choisir une structure de données qui réponde aux besoins des statisticiens.

D'autre part le fait d'envisager le logiciel adaptable à des petites machines a rendu laborieuse la programmation car il fallait considérer comme paramètre variable la dimension de matrices pour les calculs.

Le code exécutable du programme occupe approximativement 30 K mots de 36 bits de mémoire centrale et peut être adapté pour tourner sur des micros de 16 bits tels que IBM-PC-XT, KAYPRO-10 avec le compilateur "FORTRAN Microsoft version 3.20" et le système d'exploitation MS-DOS.

En ce qui concerne les possibilités d'extension du système implémenté, il est possible d'inclure des routines qui permettent de faire l'examen des résidus pour tester la validité du modèle.

En outre on pourra ajouter des routines pour faire l'estimation des données manquantes.

Des nouvelles versions interactives des programmes statistiques peuvent être implémentées en utilisant les routines de traitement des données élaborées dans ce travail.



Enfin, il faut remarquer que la méthode de régression pas à pas utilisée pour résoudre les problèmes de régression, ne donne pas toujours la solution optimale. En dépit de ceci, elle reste un outil de travail pour les statisticiens qui dans la pratique choisissent l'équation en considérant des facteurs additionnels.

## BIBLIOGRAPHIE

- (1) CHAMBERS J.M.  
Statistical Computing: history and trends.  
Amer. Statist., Vol 34, 1980.
- (2) CLAUSTRIAUX J.  
A propos des logiciels statistiques.  
Document interne.  
Gembloux, Facultés des Sciences Agronomiques, 1983.
- (3) DAGNELIE P.  
Programmes statistiques TPD (Version 3): présentation générale.  
Document interne.  
Gembloux, Facultés des Sciences Agronomiques, 1983.
- (4) DAGNELIE P.  
Théorie et méthodes statistiques (Vol. 1) .  
Gembloux, Presses Agronomiques, 1980.
- (5) DAGNELIE P.  
Théorie et méthodes statistiques (Vol. 2).  
Gembloux, Presses Agronomiques, 1980.
- (6) DAGNELIE P.  
Analyse statistique à plusieurs variables.  
Gembloux, Presses Agronomiques, 1982.
- (7) DIXON W.  
Biomedical Computer Programs.  
Berkeley, Univ. California Press.
- (8) DIGITAL EQUIPEMENT CORPORATION.  
FORTRAN Langage Manual.  
February 1983.
- (9) DRAPER N., SMITH H.  
Applied Regression Analysis.  
John Wiley, 1966.
- (10) JACKSON M.A.  
System Development.  
Prentice Hall International, 1983.
- (11) EFROYSOM M. A.  
Multiple regression analysis.  
In: RAISON A. et WILF H.  
Mathematical methods for digital computers (Vol. 1).  
New York, Wiley, 1960.



- (12) GERARD G.  
Régression Multiple.  
In: Delincé J. et Mouchart M.  
Analyse de régression.  
Université Catholique de Louvain, 1978.
- (13) GERARD G. et SCHIFFLERS E.  
Une application de l'analyse de régression en physique.  
In: Delincé J et Mouchart M.  
Analyse de régression.  
Université Catholique de Louvain, 1978.
- (14) JENNRICH R.  
Stepwise Regression.  
In: Statistical methods for digital computers.  
New York, Wiley , 1976.
- (15) WOOD J.T.  
Statistical Computing.  
European Journal of Operational Research, Vol.9, N° 4, April 1982.
- (16) SCHIFFLERS E.  
Usage de l'ordinateur en regression multiple.  
In: Delincé J. et Mouchart M.  
Analyse de régression.  
Université Catholique de Louvain, 1978.
- (17) Microsoft Fortran Compiler  
for the MS-DOS Operating Système  
User's Guide  
Microsoft Corporation, 1984.
- (18) MOSTELLER F. et TUKEY W.  
Data Analyse and Regression  
Addison-Wesley Publishing Company, Inc., 1977.
- (19) ROMANO A.  
Applied Statistics for Science and Industry.  
Allyn and Bacon, Inc., 1977.
- (20) HAINAUT J.-L.  
Résumé du Modèle d'Accès.  
Notes de cours 1984  
Institut d'Informatique, Namur.

FACULTES UNIVERSITAIRES NOTRE - DAME DE LA PAIX - NAMUR  
INSTITUT D'INFORMATIQUE

IMPLEMENTATION D'UNE VERSION INTERACTIVE  
POUR L'ANALYSE DE REGRESSION MULTIPLE  
(ANNEXES)

Mémoire présenté par

Maria Hallo

en vue de l'obtention du grade de  
Licencié et Maître en Informatique

ANNEE ACADEMIQUE 1983-1984



# INDEX

=====

DEFINITION DES VARIABLES GLOBALES.....	1
PROGRAMME PRINCIPAL.....	4
SUBROUTINE TITR.....	4
SUBROUTINE IMMESS.....	6
ROUTINES DE GESTION DES MENUS	
SUBROUTINE ANSWER.....	8
SUBROUTINE IMENU .....	9
SUBROUTINE INFOR .....	10
SUBROUTINE MENUP .....	11
SUBROUTINE OPENF .....	12
SUBROUTINE FRMENU.....	13
SUBROUTINE TITLE.....	14
SUBROUTINE VALCH.....	15
ROUTINES DE CHARGEMENT DES VARIABLES	
SUBROUTINE CHV.....	17
SUBROUTINE CHVAR.....	18
SUBROUTINE CHVARF.....	19
SUBROUTINE CHVARE.....	20
SUBROUTINE CHIND.....	21
SUBROUTINE CHINF.....	22
SUBROUTINE CHINE.....	24
SUBROUTINE EXEMP.....	25
SUBROUTINE EX.....	25
ROUTINE DE SUPPRESSION DES VARIABLES	
SUBROUTINE ELIMV.....	27
ROUTINES DE TRANSFORMATION DES VARIABLES	
SUBROUTINE TRVAR.....	29
ROUTINES D'IMPRESSION DES VARIABLES	
SUBROUTINE IMP.....	37
SUBROUTINE IMPTIT.....	38
SUBROUTINE IMPPV.....	39
SUBROUTINE IMPVAR.....	43
SUBROUTINE FMI.....	45
ROUTINES POUR ACCES AUX STRUCTURES DES DONNEES	
SUBROUTINE TYPEAR.....	47
SUBROUTINE LETPD.....	52
SUBROUTINE LIECDV.....	53
SUBROUTINE LIRTPD.....	55
ROUTINES ASSOCIEES A LA REGRESSION MULTIPLE	
SUBROUTINE RM.....	58
SUBROUTINE CCF.....	60
SUBROUTINE CCRCVE.....	61
SUBROUTINE CS1.....	62
SUBROUTINE CS2.....	63
SUBROUTINE CF.....	66



SUBROUTINE	CVS.....	67
SUBROUTINE	EVSD.....	68
SUBROUTINE	FTM.....	69
SUBROUTINE	IPVID.....	70
SUBROUTINE	ISCP.....	72
SUBROUTINE	IV.....	74
SUBROUTINE	LCHOIX.....	75
SUBROUTINE	LVDV.....	76
SUBROUTINE	LNOMF.....	78
SUBROUTINE	REDD.....	79
SUBROUTINE	REG.....	80
SUBROUTINE	TVE.....	84
SUBROUTINE	TVS.....	85



# DEFINITION DES VARIABLES GLOBALES

\*\*\*\*\* \*\*\* \*\*\*\*\*

NOM	TYPE	DEFINITION
A	REAL	MATRICE DES SOMMES DES CARRES ET DES PRODUITS DES ECARTS
B	REAL	TABLEAU DES COEFFICIENTS DE REGRESSION
COV	REAL	TABLEAU DES COVARIANCES
D	REAL	TABLEAU DE SOMMES DES CARRES DES ECARTS
DONNEE	REAL	TABLEAU DES DONNEES D'UNE VARIABLE
DEF	REAL	TABLEAU DES NUMEROS DES VARIABLES DEPENDANTES
E	REAL	TABLEAU DES RESIDUS DE LA REGRESSION
ECT	REAL	TABLEAU DES ECARTS TYPES
F	REAL	TABLEAU DES VALEURS DE F DE SNEDECOR CALCULEES POUR TESTER LA SIGNIFICATION DES COEFFICIENTS DE LA REGRESSION.
IENLT	INTEGER	TABLEAU QUI CONTIENT LES POINTEURS POUR L'ENCHAINEMENT DES ENREGISTREMENTS DISPONIBLES POUR STOCKER LES TITRES.
INIT	INTEGER	TABLEAU QUI CONTIENT LES POINTEURS POUR ENCHAINER LES ENREGISTREMENTS DES TITRES AUX ENREGISTREMENTS DES DONNEES.
IENLR	INTEGER	TABLEAU QUI CONTIENT LES POINTEURS POUR L'ENCHAINEMENT DES ENREGISTREMENTS DES DONNEES.
IND	INTEGER	TABLEAU DES NUMEROS DE VARIABLES EXPLICATIVES.
LREGT	INTEGER	NOMBRE DES CARACTERES PAR ENREGISTREMENT DU FICHIER DES TITRES.
LREGP	INTEGER	NOMBRE DES MOTS PAR ENREGISTREMENT DU FICHIER DE PARAMETRES.
LREGD	INTEGER	NOMBRE DES MOTS PAR ENREGISTREMENT DU FICHIER DES DONNEES.
M	INTEGER	NOMBRE DE VARIABLES EXPLICATIVES PLUS VARIABLES DEPENDANTES.
N	INTEGER	NOMBRE DE DONNEES D'UNE VARIABLE.
NVI	INTEGER	NOMBRE DE VARIABLES INDEPENDANTES.
NVO	INTEGER	NOMBRE DE VARIABLES OBLIGATOIRES.
NVD	INTEGER	NOMBRE DE VARIABLES DEPENDANTES.



NOM0	STRING	PARTIE COMUN DES NOMS DES FICHIERS DE TITRES PARAMETRES ET DONNEES.
NOM1	STRING	NOM DU FICHIER DE TITRES.
NOM2	STRING	NOM DU FICHIER DE PARAMETRES.
NOM3	STRING	NOM DU FICHIER DE DONNEES.
NOM4	STRING	NOM DU FICHIER DE POINTEURS.
NOM	STRING	NOM DU FICHIER DE DONNEES CREE PAR L'UTILISATEUR.
P	REAL	TABEAU QUI CONTIENT LES PARAMETRES.
R	REAL	TABEAU DE COEFFICIENTS DE CORRELATION PARTIELLE.
S	REAL	TABEAU DES ERREURS STANDARD POUR LES COEFFICIENTS DE REGRESSION.
TT	INTEGER	VARIABLE QUI INDIQUE LE SOMMET DE LA FILE DES ENREGISTREMENTS OCCUPES DES TITRES.
TTDIS	INTEGER	VARIABLE QUI INDIQUE LE SOMMET DE LA FILE DES ENREGISTREMENTS DISPONIBLES POUR LES DONNEES.
TIT	REAL	TABEAU DES TITRES D'UNE VARIABLE.
V	REAL	TABEAU DES COEFFICIENTS DE VARIATION.
XM	REAL	TABEAU DES VALEURS MOYENNES.
XMAX	REAL	TABEAU DES VALEURS MAXIMUMS.
XMIN	REAL	TABEAU DES VALEURS MINIMUMS.
X	REAL	MATRICE DE DONNEES DES VARIABLES.
Y	REAL	TABEAU DES VALEURS ESTIMEES POUR LA VARIABLE DEPENDANTE.



```

PROGRAM TEST
DATA MANS/9/, KA/1/, NA/1/, MA/1/ , TEST/.FALSE./, NCP/3/
CHARACTER*1 ANS
LOGICAL TEST
INCLUDE 'COM1.DAT'
DATA K1/20/, K2/21/, K3/22/, IN/5/, LR1/400/, NMV/100/, NENR/1000/
* LR2/20/, LR3/100/, LREGT/192/, LREGP/10/,
* LREGD/50/, K5/25/, K6/26/, K7/27/, K8/28/, K9/29/, TOL/0.01/
CALL OPENF(NA,MA)
  CALL TITR
  WRITE (IN, '//
* , 1X, 'BIENVENUE AU PROGRAMME DE REGRESSION MULTIPLE', '///')
  WRITE(IN,*)
* 'SOUSHAITEZ-VOUS DES INFORMATIONS SUR LA MANIPULATION'
  WRITE(IN, '1X, 'DE CE PROGRAMME? (O/N): ', '$')
  READ(IN, 'A1')ANS
  IF(ANS.EQ. 'O')THEN
    CALL IMMESS
  END IF
  WRITE (IN,*)
* 'DANS CHAQUE ETAPE TAPCZ I SI VOUS SOUSHAITEZ'
  WRITE(IN,*) 'CONSULTER LES OPTIONS POSSIBLES'
DO WHILE (.NOT.TEST)
  CALL MENUP
  WRITE(IN,20)
  READ(IN,10)ANS
  WRITE(IN,*)
  IF(ANS.EQ. 'N')THEN
    TEST=.TRUE.
  END IF
END DO
CLOSE(UNIT=K1)
CLOSE(UNIT=K2)
STOP
10 FORMAT(A1)
20 FORMAT(// 1X, 'VOUDRIEZ-VOUS CONTINUER L''EXECUTION'',
* 1X, 'DU PROGRAMME (O/N): ', '$')
END

SUBROUTINE TITR
INCLUDE 'COM1.DAT'
C IMPRESSION DU MESSAGE INITIAL
WRITE(IN,*)
WRITE(IN,*)
WRITE(IN,*)
WRITE(IN,*)
WRITE(IN,*)
* ' RRRRRRRRRRRR NM MM UU UU LL '
  WRITE(IN,*)
* ' RR RR MMM MMM UU UU LL '
  WRITE(IN,*)
* ' RR RR MM MM MM MM UU UU LL '
  WRITE(IN,*)
* ' RRRRRRRRRRRR MM MM MM MM UU UU LL '
  WRITE(IN,*)
* ' RR RR MM MMM MM UU UU LL '
  WRITE(IN,*)
* ' RR RR MM M MM UU UU LL '
  WRITE(IN,*)
* ' RR RR MM MM UUUUUUUUUUUU LLLLLLLLLL

```

WRITE(IN,\* )  
WRITE(IN,\* )  
WRITE(IN,\* )  
RETURN  
END



```

C*****
C
C  IMMESS : IMPRIME LES MESSAGES D'INFORMATION SUR L'UTILISATION DU
C          PROGRAMME
C
C  VARIABLES GLOBALES : IN
C
C*****

```

```

SUBROUTINE IMMESS
  INCLUDE 'COM1.DAT'
  WRITE(IN,*) '          PROGRAMME RMUL'
  WRITE(IN,*) '          ====='
  WRITE(IN,*)
  WRITE(IN,*) '  CE PROGRAMME PERMET DE: '
  WRITE(IN,*)
  WRITE(IN,*)
  * '- CREER DES FICHIERS PERMANENTS POUR STOCKER LES TITRES'
  WRITE(IN,*)
  * '  PARAMETRES ET DONNEES DES VARIABLES A UTILISER POUR'
  WRITE(IN,*)
  * '  EFFECTUER DES ANALYSES STATISTIQUES.'
  WRITE(IN,*)
  * '  CHARGER DES VARIABLES A PARTIR DU TERMINAL OU DES'
  WRITE(IN,*)
  * '  FICHIERS TEMPORELS CREEES PAR L''UTILISATEUR.'
  WRITE(IN,*)
  * '  CREER DE NOUVELLES VARIABLES QUI SONT LES RESULTATS'
  WRITE(IN,*)
  * '  DES TRANSFORMATIONS DES VARIABLES DEJA EXISTENTES.'
  WRITE(IN,*)
  * '  EFFECTUER LA REGRESSION MULTIPLE.'
  WRITE(IN,*)
  * '  ELIMINER UNE OU PLUSIEURS VARIABLES CONTENUES DANS '
  WRITE(IN,*)
  * '  LES FICHIERS PERMANENTS.'
  WRITE(IN,*)
  * '  IMPRIMER LE RESULTAT DE LA REGRESSION ET LES RESIDUS'
  WRITE(IN,*)
  * '  SUR DISQUE ET/OU SUR PAPIER.'
  WRITE(IN,*)
  * '  IMPRIMER LA OU LES VARIABLES CONTENUES DANS LES '
  WRITE(IN,*)
  * '  FICHIERS DES DONNEES PERMANENTS.'
  WRITE(IN,*)
  * '  L''EXECUTION DE LA REGRESSION MULTIPLE REQUIERT '
  WRITE(IN,*)
  * '  L''EXISTENCE D''UN FICHIER DES DONNEES PERMANENTS CREE'
  WRITE(IN,*)
  * '  PAR CE PROGRAMME LORS DU CHARGEMENT DES VARIABLES.'
  RETURN
END

```



\*\*\*\*\*

ROUTINES DE GESTION DES MENUS

\*\*\*\*\*



```

C*****
C
C  ANSWER: CETTE ROUTINE LIT ET FAIT LA VALIDATION DU CHOIX DE
C          L'UTILISATEUR DANS CHAQUE MENU.
C
C  VARIABLES GLOBALES: IN,K2
C
C  ENTREE: MA,IN,K2
C
C  SORTIE: ANS = VARIABLE QUI CONTIENT LE CHOIX DE L'UTILISATEUR
C           TEST = .TRUE. SI LE CHOIX EST VALIDE.
C           = .FALSE. DANS LE CAS CONTRAIRE.
C
C*****

```

```

      SUBROUTINE ANSWER(ANS,TEST,MA,J)
      INCLUDE 'COM1.DAT'
      LOGICAL TEST
      DIMENSION ANSW(10)
      CHARACTER*2 ANS,ANSW,AST
      DATA AST/'**'/
C      LECTURE DES CHOIX POSSIBLES
      READ(K2,MA,100)(ANSW(I),I=1,10)
C      *LECTURE AVEC FORMATO LIBRE
      WRITE(IN,'1X','*  '$')
      READ(IN,200)ANS
      I=1
      TEST=.FALSE.
      DO WHILE ((ANSW(I).NE.AST).AND.(I.LE.10))
        IF (ANSW(I).EQ.ANS) THEN
          J=I
          TEST=.TRUE.
        END IF
        I=I+1
      END DO
      RETURN
100    FORMAT(10A2)
200    FORMAT(A2)
      END

```

9

```

C*****C
C
C  IMENU: CETTE ROUTINE LIT ET AFFICHE A L'ECRAN LE MENU IDENTIFIE PAR C
C      LE NUMERO NA C
C
C  VARIABLES GLOBALES: K1,JN,MENU C
C
C  ENTREE: K1,IN,NA C
C*****C

```

```

      SUBROUTINE IMENU(NA)
      INCLUDE 'COM1.DAT'
      DIMENSION MENU(10)
      CHARACTER*40 MENU
      READ(K1,NA,1)(MENU(I),I=1,MANS)
      IMANSX = MANS - 2
      WRITE(IN,2)(MENU(I),I=1,IMANSX)
      RETURN
1      FORMAT(40X,9(A40))
2      FORMAT(/(20X,A40/))
      END

```



```

C*****C
C
C  INFOR: CETTE ROUTINE AFFICHE A L'ECRAN LE MENU IDENTIFIE PAR LE
C          NUMERO NA, SUR DEMANDE DE L'UTILISATEUR.
C
C  VARIABLES GLOBALES: IN,MANS,MENU
C
C  ENTREE: IN,MANS,ANS
C          TEST = .TRUE. SI LE CHOIX DE L'UTILISATEUR APPARTIENT A UN
C          DES CHOIX CORRECTS
C          ANS = CHOIX DE L'UTILISATEUR
C
C*****C

```

```

SUBROUTINE INFOR(NA,TEST,ANS)
INCLUDE 'COM1.DAT'
LOGICAL TEST
CHARACTER*2 INF,ANS
DATA INF/'I '/
  IF (TEST)THEN
    WRITE(IN,*)' '
    IF( ANS.EQ.INF)THEN
      WRITE(IN,*)' '
      CALL IMENU(NA)
      TEST=.FALSE.
    END IF
  END IF
RETURN
END

```

```

C*****
C
C  MENUP: CETTE ROUTINE PERMET D'AFFICHER A L'ECRAN LE MENU PRINCIPAL
C          VALIDER LES REPONSES DE L'UTILISATEUR ET EFFECTUER LE
C          TRAITEMENT CHOISI
C
C  ROUTINES APPELEES: ELIMV,TRVAR,IMP,RM
C
C*****

```

```

      SUBROUTINE MENUP
      INCLUDE 'COM1.DAT'
      IFIN=.FALSE.
      DO WHILE (.NOT.IFIN)
      MANS = 7
      CALL PMENU(1,1,J)
      IF(J.NE.6)THEN
        CALL TYPEAB(NVA,2,IERR)
      END IF
      GO TO(50,60,70,80,90,100)J
C      CHARGEMENT DES VARIABLES
50     CALL CHV
      RETURN
C      ELIMINATION DES VARIABLES
60     CALL ELIMV
      RETURN
C      TRANSFORMATION DES VARIABLES
70     CALL TRVAR
      RETURN
C      IMPRESSION DES VARIABLES
80     CALL IMP
      RETURN
90     CALL RM
      RETURN
100    IFIN=.TRUE.
      END DO
      RETURN
      END

```



```

C*****
C
C   OPENF: CETTE ROUTINE OUVRE LES FICHIERS ASSOCIES A L'AFFICHAGE
C         D'UN MENU
C
C   VARIABLES GLOBALES: K1,K2,LR1,LR2
C
C*****

```

```

      SUBROUTINE OPENF(NA,MA)
      INCLUDE 'COM1.DAT'
      OPEN(UNIT=K1,FILE='TMENU.DAT',ASSOCIATE VARIABLE=NA,ACCESS='DIR
*   ECT', CARRIAGECONTROL='LIST',RECORDSIZE=LR1,FORM='FORMATTED')
      OPEN(UNIT=K2,FILE='ANSWER.DAT',ASSOCIATE VARIABLE=MA,ACCESS='DIRE
*   CT', CARRIAGECONTROL='LIST',RECORDSIZE=LR2,FORM='FORMATTED')
      RETURN
      END

```





```
C*****C
C
C  TITLE : CETTE ROUTINE PERMET D'AFFICHER LE TITRE DU MENU IDENTIFIE C
C          PAR NA. C
C
C  VARIABLES GLOBALES: IN,K1,TITRE C
C
C  ENTREE: IN,K1 C
C
C  SORTIE: TITRE C
C*****C
```

```
      SUBROUTINE TITLE(NA)
      INCLUDE 'COM1.DAT'
      CHARACTER*40 TITRE
      READ(K1,NA,1)TITRE
      WRITE(IN,2)TITRE
      RETURN
1      FORMAT(A40)
2      FORMAT(/20X,A40/)
      END
```

```

C*****C
C
C VALCH:   CETTE ROUTINE LIT ET FAIT LA VALIDATION DU CHOIX ET
C          AFFICHE LE MENU SUR DEMANDE DE L'UTILISATEUR
C
C VARIABLES GLOBALES: MANS,IN,K1,K2
C
C ENTREE: MANS,IN,K1,K2,NA,MA
C
C SORTIE: J = NUMERO QUI IDENTIFIE LE CHOIX DE L'UTILISATEUR DANS
C          CHAQUE MENU
C
C*****C

```

```

SUBROUTINE VALCH(NA,MA,J)
INCLUDE 'COM1.DAT'
LOGICAL TEST
CHARACTER*2 ANS
TEST=.FALSE.
DO WHILE (.NOT.TEST)
  CALL ANSWER(ANS,TEST,MA,J)
  IF (.NOT. TEST) THEN
    WRITE(IN,*)'    '
  END IF
  CALL INFOR(NA,TEST,ANS)
END DO
RETURN
END

```



\*\*\*\*\*  
ROUTINES DE CHARGEMENT DES VARIABLES  
\*\*\*\*\*

```

C*****C
C
C   CHV: CETTE ROUTINE PRESENTE LE MENU DE CHARGEMENT DES VARIABLES ET C
C       EFFECTUE LE TRAITEMENT CHOISI PAR L'UTILISATEUR C
C C
C   ROUTINES APPELEES: CHVAR,CHIND C
C C
C*****C

```

```

      SUBROUTINE CHV
      LOGICAL TEST
      INCLUDE 'COM1.DAT'
      MANS = 5
      CALL PRMENU(2,2,J)
      GO TO (140,150)J
C   CHARGEMENT VARIABLE PAR VARIABLE
140  CALL CHVAR
      RETURN
C   CHARGEMENT INDIVIDU PAR INDIVIDU
150  CALL CHIND
      RETURN
      END

```



```

C*****
C
C      CHVAR: CETTE ROUTINE EFFECTUE LE CHARGEMENT DES DONNEES VARIABLE
C          PAR VARIABLE
C
C      VARIABLE GLOBAL: IN
C
C*****

```

```

      SUBROUTINE CHVAR
      CHARACTER * 1 CHOIX
      INCLUDE 'COM1.DAT'
      WRITE(IN,5)
      READ(IN,'A1')CHOIX
      IF (CHOIX.EQ.'F')THEN
        CALL CHVARF
      ELSE
        CALL CHVARE
      END IF
      RETURN
5      FORMAT(//20X,'MANIERE D''INTRODUIRE LES DONNEES'///,25X,
* 'F : PAR FICHIER' ///,25X, 'E : PAR ECRAN'///)
      END

```

```

C*****
C
C CHVARF: CETTE ROUTINE PERMET DE LIRE LES TITRES,PARAMETRES ET DONNEES
C          D'UNE OU PLUSIEURS VARIABLES A PARTIR D'UN FICHIER DES DONNEES
C          TEMPORAIRES CREE PAR L'UTILISATEUR ET D'AJOUTER AUX FICHIERS
C          DES VARIABLES
C
C ROUTINE APPELEE: TYPEAB
C VARIABLES GLOBALES: IN,TJT,PAR,DONNEE,K8
C
C*****

```

```

SUBROUTINE CHVARF
INCLUDE 'COM1.DAT'
CHARACTER*48 BLANC
DATA BLANC/'
LOGICAL TEST
CHARACTER*1 CHOIX
WRITE(IN,*)'SOUSHAITEZ-VOUS DES INFORMATIONS SUR LE FICHIER'
WRITE(IN,'1X','A CREER PAR L UTILISATEUR ? (O/N): ','$')
READ(IN,'A1')CHOIX
IF(CHOIX.EQ.'O')THEN
    CALL EXEMP
    RETURN
END IF
CALL LNMFF(TEST)
IF (TEST) THEN
    RETURN
END IF
WRITE(IN,'1X','NOMBRE DES VARIABLES A CHARGER: ','$')
READ(IN,*)NVCH
C LECTURE DES TITRES,PARAMETRES ET DONNEES A PARTIR D'UN FICHIER
C SEQUENTIEL ET ECRITURE SUR DISQUE
DO IK=1,NVCH
    DO I=1,4
        TIT(I)=BLANC
    END DO
    DO I=1,4
        READ(K8,'A48',ERR=10)TIT(I)
    END DO
    READ(K8,*,ERR=20)(P(I),I=1,10)
    READ(K8,*,ERR=30)ND
    DO I=1,1000
        DONNEE(I)=0
    END DO
    READ(K8,*,ERR=40)(DONNEE(J),J=1,ND)
    CALL TYPEAB(NVA,1,IERR)
END DO
CLOSE(K8)
RETURN
10 WRITE(IN,*)'ERREUR EN LECTURE DES TITRES'
RETURN
20 WRITE(IN,*)'ERREUR EN LECTURE DES PARAMETRES'
RETURN
30 WRITE(IN,*)'ERREUR EN LECTURE DE NOMBRE DES DONNEES'
RETURN
40 WRITE(IN,*)'ERREUR EN LECTURE DES DONNEES'
RETURN
END

```



```

C*****
C
C  CHVARE:  CETTE ROUTINE EFFECTUE LE CHARGEMENT DES DONNEES
C           VARIABLE PAR VARIABLE A PARTIR DU TERMINAL
C
C  VARIABLES GLOBALES:  IN,TIT,P,DONNEE
C
C  ENTREE:  IN
C
C*****

```

```

SUBROUTINE CHVARE
INCLUDE 'COM1.DAT'
WRITE(IN,*)'                                CHARGEMENT VARIABLE PAR VARIABLE'
WRITE(IN,*)'                                ====='
WRITE(IN,*)
WRITE(IN,*)'INTRODUCTION DES TITRES'
WRITE(IN,*)
WRITE(IN,*)'NOM DE L.' 'EXPERIMENTATEUR: '
READ(IN,'A48')TIT(1)
WRITE(IN,*)'NOM DE L.' 'EXPERIENCE:'
READ(IN,'A48')TIT(2)
WRITE(IN,*)'NOM DE LA VARIABLE:'
READ(IN,'A48')TIT(3)
WRITE(IN,*)'TYPE DE TRANSFORMATION EFFECTUEE SUR LA VARIABLE:'
READ(IN,'A48')TIT(4)
DO I = 1,4
  P(I) = 0
END DO
WRITE(IN,*)'DONNEZ LE CODE NUMERIQUE DE FIN DE DONNEES:'
READ(IN,*)P(8)
WRITE(IN,*)'DONNEZ LE CODE NUMERIQUE DE DONNEES MANQUANTES:'
READ(IN,*)P(9)
WRITE(IN,'1X','NOMBRE DE DONNEES A CHARGER:','',$')
READ(IN,*)ND
P(6)=ND
WRITE(IN,*)'DONNEES (FORMAT LIBRE):'
READ(IN,*)(DONNEE(I),I=1,ND)
CALL TYPEAB(NVA,1,IERR)
CALL TYPEAB(NVA,4,IERR)
RETURN
END

```

C\*\*\*\*\*

C  
C CHIND: CETTE ROUTINE EFFECTUE LE CHARGEMENT DES DONNEES INDIVIDU  
C PAR INDIVIDU

C  
C VARIABLE GLOBAL: IN  
C

C\*\*\*\*\*

```
SUBROUTINE CHIND
CHARACTER * 1 CHOIX
INCLUDE 'COM1.DAT'
WRITE(IN,5)
READ(IN,'A1')CHOIX
  IF (CHOIX.EQ.'F')THEN
    CALL CHINF
  ELSE
    CALL CHINE
  END IF
RETURN
```

```
5  FORMAT(//20X,'MANIERE D''INTRODUIRE LES DONNEES'///,25X,
*  'F : PAR FICHIER' //,25X, 'E : PAR ECRAN'//)
END
```

@



```

C*****
C
C CHINF: CETTE ROUTINE PERMET DE LIRE LES TITRES,PARAMETRES ET DONNEES
C      DES VARIABLES CODES INDIVIDU PAR INDIVIDU A PARTIR DES FICHIERS
C      TEMPORAIRES CREE PAR L'UTILISATEUR ET D'AJOUTER AUX FICHIERS
C      DES VARIABLES
C
C ROUTINE APPELEE: TYPEAB
C VARIABLES GLOBALES: IN,TIT,PAR,DONNEE,K8
C
C*****

```

```

SUBROUTINE CHINF
  INCLUDE 'COM1.DAT'
  CHARACTER*48 BLANC
  DATA BLANC/'
  LOGICAL TEST
  CHARACTER*1 CHOIX
  WRITE(IN,*)'SOUHAITEZ-VOUS DES INFORMATIONS SUR LE FICHIER'
  WRITE(IN,'1X','A CREER PAR L UTILISATEUR ? (O/N): ','$')
  READ(IN,'A1')CHOIX
  IF(CHOIX.EQ.'O')THEN
    CALL EX
    RETURN
  END IF
  CALL LNOMF(TEST)
  IF (TEST) THEN
    RETURN
  END IF
  DO I=1,4
    TIT(I)=BLANC
  END DO
  DO I=1,2
    READ(K8,'A48',ERR=20)TIT(I)
  END DO
C  LECTURE DU NOMBRE DES INDIVIDUS ET DU NOMBRE DES VARIABLES
C  PAR INDIVIDU
  READ(K8,*,ERR=10)NIND,NVPI
  READ(K8,*,ERR=30)(P(I),I=1,10)
  DO I=1,NIND
    READ(K8,*,ERR=40)(A(I,J),J=1,NVPI)
  END DO
  DO J=1,NVPI
    DO I=1,1000
      DONNEE(I)=0.
    END DO
    DO I=1,NIND
      DONNEE(I)=A(I,J)
    END DO
    DO I=2,3
      TIT(I)=BLANC
      READ(K8,*,ERR=40)TIT(I)
    END DO
    CALL TYPEAB(NVA,1,IERR)
  END DO
  CLOSE(K8)
  RETURN
10  WRITE(IN,*)'ERREUR EN LECTURE DU NOMBRE DES INDIVIDUS'
    RETURN
20  WRITE(IN,*)'ERREUR EN LECTURE DE TITRES GENERAUX'

```



```
      RETURN  
30  WRITE(IN,*) 'ERREUR EN LECTURE DE PARAMETRES'  
      RETURN  
40  WRITE(IN,*) 'ERREUR EN LECTURE DES DONNEES'  
      RETURN  
      END
```



```

C*****
C
C  CHINE:  CETTE ROUTINE EFFECTUE LE CHARGEMENT DES DONNEES
C          INDIVIDU PAR INDIVIDU A PARTIR DU TERMINAL
C
C  VARIABLES GLOBALES: IN,TIT,P,DONNEE
C
C  ENTREE: IN
C
C*****

```

```

SUBROUTINE CHINE
  INCLUDE 'COM1.DAT'
  CHARACTER*48 BLANC
  DATA BLANC/'
  WRITE(IN,*)
  WRITE(IN,*) 'CHARGEMENT INDIVIDU PAR INDIVIDU'
  WRITE(IN,*)
  WRITE(IN,*) 'INTRODUCTION DES TITRES GENERAUX'
  WRITE(IN,*)
  WRITE(IN,*) 'NOM DE L''EXPERIMENTATEUR:'
  READ(IN,'A48') TIT(1)
  WRITE(IN,*) 'NOM DE L''EXPERIENCE:'
  READ(IN,'A48') TIT(2)
  WRITE(IN,1)
1  FORMAT(1X,'NOMBRE D''INDIVIDUS:',$,)
  READ(IN,*) NIND
  P(5)=NIND
  WRITE(IN,'1X,'NOMBRE DE VARIABLES PAR INDIVIDU:',$,)
  READ(IN,*) NVPI
  DO I=1,10
    P(I)=0.
  END DO
  WRITE(IN,'1X,'CODE DE FIN DES DONNEES:',$,)
  READ(IN,*) P(8)
  WRITE(IN,'1X,'CODE NUMERIQUE DES DONNEES MANQUANTES:',$,)
  READ(IN,*) P(9)
  WRITE(IN,*) 'INTRODUCTION DES DONNEES'
  WRITE(IN,*) 'UNE LIGNE PAR INDIVIDU:'
  DO I=1,NIND
    READ(IN,*)(A(I,J),J=1,NVPI)
  END DO
  WRITE(IN,*) 'INTRODUCTION DES TITRES POUR CHAQUE VARIABLE'
  DO J=1,NVPI
    DO I=1,NIND
      DONNEE(I)=A(I,J)
    END DO
    TIT(1)=BLANC
    WRITE(IN,*) 'NOM DE LA VARIABLE:'
    READ(IN,*) TIT(1)
    TIT(2)=BLANC
    WRITE(IN,*) 'TYPE DE TRANSFORMATION:'
    READ(IN,*) TIT(2)
    CALL TYPEAB(NVA,1,IERR)
  END DO
  RETURN
END

```

```

C*****
C
C   EXEMP : CETTE ROUTINE PERMET D'AFFICHER A L'ECRAN UN EXEMPLE DE
C           FICHER DES DONNEES CODEES VARIABLE PAR VARIABLE
C
C*****

```

```

      SUBROUTINE EXEMP
      INCLUDE 'COM1.DAT'
      CHARACTER*80 LINEA
      K10=30
      OPEN(UNIT=K10,NAME='EXREG.DAT',ACCESS='SEQIN')
10    READ(K10,'A80',END=20)LINEA
      WRITE(IN,'A80')LINEA
      GO TO 10
20    IFIN=.TRUE.
      CLOSE(UNIT=K10)
      RETURN
      END

```

```

C*****
C
C   EX :  CETTE ROUTINE PERMET D'AFFICHER A L'ECRAN UN EXEMPLE DE
C           FICHER DES DONNEES CODEES INDIVIDU PAR INDIVIDU
C
C*****

```

```

      SUBROUTINE EX
      INCLUDE 'COM1.DAT'
      CHARACTER*80 LINEA
      K11=31
      WRITE(IN,*)
      OPEN(UNIT=K11,NAME='EXIND.DAT',ACCESS='SEQIN')
10    READ(K11,'A80',END=20)LINEA
      WRITE(IN,'A80')LINEA
      GO TO 10
20    IFIN=.TRUE.
      CLOSE(UNIT=K11)
      RETURN
      END

```



\*\*\*\*\*

ROUTINES DE SUPPRESSION DES VARIABLES

\*\*\*\*\*

```

*****
C
C ELIMV: CETTE ROUTINE ELIMINE UNE OU PLUSIEURS VARIABLES A
C PARTIR DES FICHIERS DES VARIABLES
C
C ROUTINE APPELEE: TYPEAB
C VARIABLES GLOBALES: IN,TIT
C
*****

```

```

SUBROUTINE ELIMV
INCLUDE 'COM1.DAT'
CHARACTER*1 CHOIX
WRITE(IN,*)'ELIMINATION DES VARIABLES'
WRITE(IN,*)
WRITE(IN,'1X,
* 'VOUDRIEZ-VOUS EFFACER TOUTES LES VARIABLES ? (O/N): ','$')
READ(IN,'A1')CHOIX
IF(CHOIX.EQ.'O')THEN
    CALL TYPEAB(NVA,6,IERR)
    RETURN
END IF
WRITE(IN,*)
WRITE
*(IN,'/1X,'DONNEZ LE NOMBRE DES VARIABLES A SUPPRIMER: ','$')
READ(IN,*)NEL
IF(NEL.GT.0)THEN
    DO III=1,NEL
        CHOIX = 'N'
        I=1
        DO WHILE((CHOIX.NE.'O').AND.(I.LE.3))
            WRITE(IN,*)'DONNEZ LE NUMERO DE LA VARIABLE A SUPPRIMER'
            READ(IN,*)NVA
            CALL TYPEAB(NVA,7,IERR)
            IF(IERR.NE.1)THEN
                WRITE(IN,*)
                WRITE(IN,*)'TITRE:'
                WRITE(IN,*)
                DO IK=1,4
                    WRITE(IN,'1X,A48')TIT(IK)
                END DO
                WRITE(IN,5)
                READ(IN,'A1')CHOIX
            END IF
            I=I+1
        END DO
        IF(I.LE.3)THEN
            ELIMINATION DE LA VARIABLE
            CALL TYPEAB(NVA,5,IERR)
        END IF
    END DO
END IF
RETURN
5 FORMAT(/1X,'C'EST BIEN CETTE VARIABLE QUE',/
* 1X,'VOUS VOULEZ SUPPRIMER? (O/N): ','$')
END

```



\*\*\*\*\*  
ROUTINES DE TRANSFORMATION DES VARIABLES  
\*\*\*\*\*

```

C*****C
C
C   TRVAR : CETTE ROUTINE PERMET D'EFFECTUER DES TRANSFORMATIONS . C
C           DE VARIABLES, LES RESULTATS ETANT EUX-MEMES TRANSFERES C
C           SUR DISQUE COMME DES NOUVELLES VARIABLES C
C
C   VARIABLES GLOBALES: IN, DONNEE, TIT, P, DON C
C
C   ROUTINE APPELEE: TYPEAB C
C
C*****C

```

```

SUBROUTINE TRVAR
  INCLUDE 'COM1.DAT'
  LOGICAL TEST
  CHARACTER*4 FT, FONC(8)
  CHARACTER* 1 NCC
  DATA (FONC(I), I=1,8) / 'ABS ', 'SQRT', 'LOGE', 'LOGD', 'EXP ',
* 'NEGA', 'ZERO', 'NEZE' /
  WRITE(IN,*)
* 'DONNER LE CODE DU TYPE DE TRANSFORMATION A EFFECTUER:'
  WRITE(IN, '/15X, 'POSSIBILITES: ', ')
  WRITE(IN, 1)
1  FORMAT(15X, '1.  X2 = C * X1' / 15X, '2.  X3 = X1 + X2' /
* 15X, '3.  X3 = X1 - X2' / 15X, '4.  X3 = X1 * X2' /
* 15X, '5.  X3 = X1 / X2' / 15X, '6.  X2 = X1 ** C' /
* 15X, '7.  X3 = F(X1)' / 15X, 'FONCTIONS F DISPONIBLES: ' /
* 15X, 'ABS = VALEUR ABSOLUE' / 15X, 'SQRT= RACINE CARRE' /,
* 15X, 'LOGE = LOGARITHME NEPERIEN' /,
* 15X, 'LOGD = LOGARITHME DECIMAL' / 15X, 'EXP = EXPONENTIELLE' /,
* 15X, 'NEGA = SUPPRESSION DES VALEURS NEGATIVES' /,
* 15X, 'ZERO = SUPPRESSION DES VALEURS NULLES' /,
* 15X, 'NEZE = SUPPRESSION DES VALEURS NULLES ET NEGATIVES' /)
5242 WRITE(IN, '1X, 'CHOIX: ', '$')
  READ(IN, 'A1') NCC
  NC=8
  IF(NCC.EQ.'1') NC = 1
  IF(NCC.EQ.'2') NC = 2
  IF(NCC.EQ.'3') NC = 3
  IF(NCC.EQ.'4') NC = 4
  IF(NCC.EQ.'5') NC = 5
  IF(NCC.EQ.'6') NC = 6
  IF(NCC.EQ.'7') NC = 7
  IF(NC.EQ.8) THEN
    WRITE(IN,*) 'ERREUR ! L'OPTION DOIT ETRE NUMERIQUE'
    GO TO 5242
  END IF
  IF((NC.EQ.1).OR.(NC.EQ.6)) THEN
    WRITE(IN,*)
    WRITE(IN,*) 'DONNER LA VALEUR DE LA CONSTANCE C:'
    READ(IN,*) C
  END IF
  I=1
  TEST=.TRUE.
  DO WHILE ((TEST).AND.(I.LE.2))
    TEST = .FALSE.
    WRITE(IN, '1X, 'NUMERO DE LA VARIABLE X1: ', '$')
    READ(IN,*) NX1
    LECTURE DE LA VARIABLE NX1
    CALL TYPEAB(NX1,7,ICRR)

```



```
IF(IERR,EQ,1)THEN
  I=I+1
  TEST = .TRUE.
END IF
END DO
IF (I,GT,2)THEN
  RETURN
END IF
```

```

ND = P(6)
DO I = 1,ND
  DON(I) = DONNEE(I)
END DO
IF((NC,EQ,2),OR,(NC,EQ,3),OR,(NC,EQ,4),OR,(NC,EQ,5))THEN
  I=1
  TEST = .TRUE.
  DO WHILE((TEST),AND,(I,LE,2))
    TEST =.FALSE.
    WRITE(IN,'1X','NUMERO DE LA VARIABLE X2: ','$')
    READ(IN,*)NX2
    CALL TYPEAB(NX2,7,IERR)
    IF (IERR,EQ,1)THEN
      I=I+1
      TEST = .TRUE.
    END IF
  END DO
  IF(I,GT,2)THEN
    RETURN
  END IF
END IF
WRITE(IN,*)
WRITE(IN,'/1X','DONNEZ LES TITRES DE LA NOUVELLE VARIABLE','/ ')
WRITE(IN,*)'NOM DE L''EXPERIMENTATEUR:'
READ(IN,'A48')TIT(1)
WRITE(IN,*)'NOM DE L''EXPERIENCE:'
READ(IN,'A48')TIT(2)
WRITE(IN,*)'NOM DE LA VARIABLE:'
READ(IN,'A48')TIT(3)
WRITE(IN,*)'TYPE DE TRANSFORMATION EFFECTUEE SUR LA VARIABLE:'
READ(IN,'A48')TIT(4)
95  FORMAT(A4)
NDPV = P(6) - 1
GO TO (10,20,30,40,50,60,70)NC

C      X2 = C * X1

10     DO I = 1,NDPV
      IF (DON(I),NE,P(9))THEN
        DONNEE(I) = C * DON(I)
      END IF
    END DO
    DONNEE(NDPV+1)=P(8)
    CALL TYPEAB(NVA,1,IERR)
    CALL TYPEAB(NVA,4,IERR)
    RETURN

C      X3 = X1 + X2
20     DO I = 1,NDPV
      IF ((DONNEE(I),NE,P(9)),AND,(DON(I),NE,P(9))) THEN
        DONNEE(I) = DONNEE(I) + DON(I)
      ELSE
        DONNEE(I) = P(9)
      END IF
    END DO
    DONNEE(NDPV+1)=P(8)
    CALL TYPEAB(NVA,1,IERR)
    CALL TYPEAB(NVA,4,IERR)
    RETURN

```



```

C      X3 = X2 - X1
30    DO I = 1,NDPV
      IF((DONNEE(I),NE,P(9)),AND,(DON(I),NE,P(9))) THEN
        DONNEE(I) = DONNEE(I) - DON(I)
      ELSE
        DONNEE(I) = P(9)
      END IF
    END DO
    DONNEE(NDPV+1)=P(8)
    CALL TYPEAB(NVA,1,IERR)
    CALL TYPEAB(NVA,4,IERR)
    RETURN

```

```

C      X3 = X1 * X2
40    DO I= 1,NDPV
      IF((DONNEE(I),NE,P(9)),AND,(DON(I),NE,P(9))) THEN
        DONNEE(I) = DON(I)*DONNEE(I)
      ELSE
        DONNEE(I) = P(9)
      END IF
    END DO
    DONNEE(NDPV+1)=P(8)
    CALL TYPEAB(NVA,1,IERR)
    CALL TYPEAB(NVA,4,IERR)
    RETURN

```

```

C      X3 = X1 / X2
50    DO I= 1,NDPV
      IF((DONNEE(I),NE,P(9)),AND,(DON(I),NE,P(9))) THEN
        DONNEE(I) = DON(I)/DONNEE(I)
      ELSE
        DONNEE(I) = P(9)
      END IF
    END DO
    DONNEE(NDPV+1)=P(8)
    CALL TYPEAB(NVA,1,IERR)
    CALL TYPEAB(NVA,4,IERR)
    RETURN

```

```

C      X2 = X1 ** C
60    DO I= 1,NDPV
      IF(DON(I),NE,P(9)) THEN
        DONNEE(I) = DON(I)**C
      ELSE
        DONNEE(I) = P(9)
      END IF
    END DO
    DONNEE(NDPV+1)=P(8)
    CALL TYPEAB(NVA,1,IERR)
    CALL TYPEAB(NVA,4,IERR)
    RETURN

```

```

C      X3 = F(X1)
70    I = 1
      TEST = .TRUE.
      DO WHILE ((TEST).AND.(I.LE.2))
        TEST = .FALSE.
        WRITE(IN,*)
        WRITE(IN,'1X','INDIQUER LE NOM DE LA FONCTION F: ','$')
        READ(IN,95)FT
        J=0
        DO K = 1,8
          IF ( FT,EQ,FONC(K)) THEN
            J = K
          END IF
        END DO
        IF(J,EQ,0)THEN
          I=I+1
          TEST = .TRUE.
        END IF
        IF(I,GT,2)THEN
          RETURN
        END IF
      END DO
      GO TO (110,120,130,140,150,160,170,180) J

```

```

C      X3 = ABS(X1)
110   DO I = 1,NDFV
      IF ( DON(I),NE,P(9)) THEN
        DONNEE(I) = ABS (DON(I))
      ELSE
        DONNEE(I) = P(9)
      END IF
    END DO
    DONNEE(K+1)=P(8)
    CALL TYPEAB(NVA,1,IERR)
    CALL TYPEAB(NVA,4,IERR)
    RETURN

```

```

C      X3 =SQRT(X1)
120   DO I = 1,NDFV
      IF (( DON(I),NE,P(9)).AND.(DON(I),GT,0)) THEN
        DONNEE(I) = SQRT (DON(I))
      ELSE
        DONNEE(I) = P(9)
      END IF
    END DO
    DONNEE(NDFV+1)=P(8)
    CALL TYPEAB(NVA,1,IERR)
    CALL TYPEAB(NVA,4,IERR)
    RETURN

```



```

C      X3 = LOGE(X1)
130    DO I = 1,NDPV
        IF ( DON(I),NE,P(9)) THEN
            DONNEE(I) = ALOG (DON(I))
        ELSE
            DONNEE(I) = P(9)
        END IF
    END DO
    DONNEE(NDPV+1)=P(8)
    CALL TYPEAB(NVA,1,IERR)
    CALL TYPEAB(NVA,4,IERR)

    RETURN

```

```

C      X3 = LOGD(X1)
140    DO I = 1,NDPV
        IF ( DON(I),NE,P(9)) THEN
            DONNEE(I) = ALOG10 (DON(I))
        ELSE
            DONNEE(I) = P(9)
        END IF
    END DO
    DONNEE(NDPV+1)=P(8)
    CALL TYPEAB(NVA,1,IERR)
    CALL TYPEAB(NVA,4,IERR)
    RETURN

```

```

C      X3 = EXP(X1)
150    DO I = 1,NDPV
        IF ( DON(I),NE,P(9)) THEN
            DONNEE(I) = EXP(DON(I))
        ELSE
            DONNEE(I) = P(9)
        END IF
    END DO
    DONNEE(NDPV+1)=P(8)
    CALL TYPEAB(NVA,1,IERR)
    CALL TYPEAB(NVA,4,IERR)
    RETURN

```

```

C      SUPPRESSION DES VALEURS NEGATIVES
160    K = 0
    DO I = 1,NDPV
        IF(DON(I),GT,0)THEN
            K = K + 1
            DONNEE(K) = DON(I)
        END IF
    END DO
    DONNEE(K+1)=P(8)
    CALL TYPEAB(NVA,1,IERR)
    CALL TYPEAB(NVA,4,IERR)
    RETURN

```

C SUPPRESSION DES VALEURS NULLES

170 K = 0  
DO I = 1,NDPV  
IF(DON(I),EQ,0)THEN  
K=K+1  
DONNEE(K) = DON(I)  
END IF  
END DO  
DONNEE(K+1)=P(8)  
CALL TYPEAB(NVA,1,IERR)  
CALL TYPEAB(NVA,4,IERR)  
RETURN

C SUPPRESSION DES VALEURS NULLES ET NEGATIVES

180 K = 0  
DO I = 1,NDPV  
IF(DON(I),LE,0)THEN  
K=K+1  
DONNEE(K)=DON(I)  
END IF  
END DO  
DONNEE(K+1)=P(8)  
CALL TYPEAB(NVA,1,IERR)  
CALL TYPEAB(NVA,4,IERR)  
RETURN  
END

@



\*\*\*\*\*

IMPRESSION DES VARIABLES

\*\*\*\*\*

```

C*****
C
C  IMP: CETTE ROUTINE AFFICHE A L'ECRAN LE MENU D'IMPRESSION SUR DEMANDE
C      DE L'UTILISATEUR ET EFFECTUE LE TRAITEMENT CHOISI
C
C  VARIABLE GLOBALE: INIT
C
C  ROUTINES APPELEES: IMPTIT,IMPVAR,IMPPV
C
C*****

```

```

      SUBROUTINE IMP
      INCLUDE 'COM1.DAT'
      CHARACTER*1 REP
C      WRITE(IN,*)'IMPRESSION DES DONNEES'
      MANS = 6
      CALL PRMENU(3,3,J)
      WRITE(IN,*)'VOUDRIEZ-VOUS L''IMPRESSION A L''ECRAN '
      WRITE(IN,'1X','OU SUR PAPIER(P/E): ','$')
      READ(IN,'A1')REP
      IF (REP.EQ.'P')THEN
         IOUT=NCP
      ELSE
         IOUT=IN
      END IF
      READ(K9'2')(INIT(I),I=1,NMV)
      GO TO (90,95,100,105)J
C      IMPRESSION DES TITRES
90      CALL IMPTIT
      RETURN
C      IMPRESSION VARIABLE PAR VARIABLE
95      CALL IMPVAR
      RETURN
C      IMPRESSION DE PLUSIEURS VARIABLES
100     CALL IMPPV
      RETURN
105     IFIN=.TRUE.
      RETURN
      END

```



```

C*****C
C
C  IMPTIT: CETTE ROUTINE EFFECTUE L'IMPRESSION DES TITRES DE TOUTES
C          LES VARIABLES
C
C  VARIABLES GLOBALES IOUT,NMV,TIT,IN,
C
C  ROUTINE APPELEE: TYPEAB
C
C  ENTREE: IOUT,NMV,IN
C
C*****C

```

```

      SUBROUTINE IMPTIT
      INCLUDE 'COM1.DAT'
C  IMPRESSION DES TITRES DE TOUTES LES VARIABLES
      WRITE(IOUT,1)
      WRITE(IOUT,5)
      WRITE(IOUT,10)
      DO I=1,NMV
         IDR = I
         CALL TYPEAB(IDR,9,IERR)
         IF(IERR.EQ.1) THEN
            WRITE(IOUT,15)I,(TIT(K),K=1,4)
         END IF
      END DO
      CALL TYPEAB(NVA,4,IERR)
      RETURN
1  FORMAT('1')
5  FORMAT(19X,34('*'),/,19X,
*  '* CONTENU DES FICHIERS *',
*  /,19X,34('*')/)
10 FORMAT(17X,'# VAR',18X,'TITRE')
15 FORMAT(18X,I3,2X,A48/3(23X,A48/))
      END

```

```

C*****
C
C  IMPPV: CETTE ROUTINE IMPRIME SIMULTANEMENT PLUSIEURS VARIABLES.
C          CONTENUES DANS LES FICHIERS DE VARIABLES
C
C  VARIABLES GLOBALPES: IN,IOUT,TIT,P,DONNEE,IND,IDEP,VARIA,XMIN,XMAX,
C                      XM, ECT
C  ENTREE: IOUT,NMV
C*****

      SUBROUTINE IMPPV
      FORMAT STANDART D'IMPRESSION: /1X,I4,2X,8F10.2
      C  NOMBRE MAXIMUM DES VARIABLES A IMPRIMER AVEC LE FORMAT
      STANDART: 8
      INCLUDE 'COM1.DAT'
      LOGICAL TES
      CHARACTER*1 CHOIX ,FVAR*80
      TES=.TRUE.
      DO WHILE (TES)
        WRITE(IN,'1X','NOMBRE DE VARIABLES A IMPRIMER: ','$')
        READ(IN,*)NV
        IF((NV.GT.0).AND.(NV.LE.NMV)) THEN
          TES = .FALSE.
        ELSE
          WRITE(IN,*)'ERREUR EN NOMBRE DE VARIABLES'
        END IF
      END DO
      C  LECTURE ET VALIDATION DES NUMEROS DES VARIABLES
      K=1
      IS=1
      DO WHILE((K.LE.3).AND.(IS.EQ.1))
        IS=0
        WRITE(IN,*)'NUMEROS DES VARIABLES:'
        READ(IN,*)(IND(I),I=1,NV)
        DO IV=1,NV
          IF(INIT(IND(IV)).EQ.0) THEN
            WRITE(IN,15)IND(I)
            IS=1
          END IF
        END DO
        K=K+1
      END DO
      IF(K.GT.3)THEN
        RETURN
      END IF
      WRITE(IOUT,1)
      WRITE(IOUT,10)
      C  LECTURE E IMPRESSION DU TITRE GENERAL
      NVAR=IND(1)
      CALL TYPEAB(NVAR,9,IERR)
      WRITE(IOUT,20)(TIT(IJ),IJ=1,2)
      C  LECTURE E IMPRESSION DES TITRES DE VARIABLES
      DO J=1,NV
        NVAR=IND(J)
        CALL TYPEAB(NVAR,9,IERR)
        WRITE(IOUT,30)NVAR,(TIT(IJ),IJ=3,4)
      END DO
      C  LECTURE E IMPRESSION DES PARAMETRES
      READ(K6'IND(1))P

```



```
WRITE(IOUT,*)'PARAMETRES:'  
DN=P(6)-1  
WRITE(IOUT,40)DN,P(8),P(9)
```

```

WRITE(IN,5)
READ(IN,'A1')CHOIX
IF(CHOIX.EQ.'O')THEN
    WRITE(IN,*)'DONNER LE FORMAT D''IMPRESSION'
    READ(IN,'A80')FVAR
ELSE
    WRITE(IOUT,2)(IND(I),I=1,NV)
END IF
C LECTURE E IMPRESSION DES DONNEES
NMIN=P(6)
IXL=0
NE=0
IL=0
ILL=0
DO WHILE(IL.LT.NMIN)
    CALL FMI(NMIN,NV,IERR,NE,ILL,IL)
    IF(IERR.EQ.1) THEN
        RETURN
    END IF
    II=1
    J=1

    DO WHILE((II.LE.50).AND.(X(II,1).LT.F(8)))
        IXL=IXL+1
        IF(CHOIX.EQ.'N')THEN
            WRITE(IOUT,50)IXL,(X(II,J),J=1,NV)
        ELSE
            WRITE(IOUT,FVAR)IXL,(X(II,J),J=1,NV)
        END IF
        II=II+1
    END DO
    ILL=IL
END DO
C REDUCTION DES DONNEES
DO I=1,NV
    CALL TYPEAB(IND(I),7,IERR)
    IDEP(I)=P(6)-1
    CALL REDD(IDEP(I),XM(I),XMIN(I),XMAX(I),ECT(I),VARIA(I))
END DO
1 FORMAT('1')
WRITE(IOUT,2)(IND(I),I=1,NV)
2 FORMAT(1X,'NUM',3X,8I10/)
WRITE(IOUT,4)(IDEP(I),I=1,NV)
4 FORMAT(1X,'EFF',3X,8I10/)
WRITE(IOUT,6)(XMIN(I),I=1,NV)
6 FORMAT(1X,'MIN',3X,8F10.2/)
WRITE(IOUT,8)(XMAX(I),I=1,NV)
8 FORMAT(1X,'MAX',3X,8F10.2/)
WRITE(IOUT,12)(XM(I),I=1,NV)
12 FORMAT(1X,'MOY',3X,8F10.2/)
WRITE(IOUT,14)(ECT(I),I=1,NV)
14 FORMAT(1X,'ECT',3X,8F10.2/)
WRITE(IOUT,16)(VARIA(I),I=1,NV)
16 FORMAT(1X,'CV ',3X,8F10.2/)
RETURN
5 FORMAT(1X,'VOUDRIEZ-VOUS INDICHER UN FORMAT D''IMPRESSION'/
* 1X,'POUR LES DONNEES (O/N): ',5)
10 FORMAT(13X,54('*'),/,13X,* IMPRESSION SIMULTANEE DE ',
* 'PLUSIEURS VARIABLES *',/,13X, 54('*'),/)
15 FORMAT(1X,'LA VARIABLE #:',I4,' N''EXISTE PAS DANS LE',

```



42

```
* ' FICHER' )
20  FORMAT(/1X,A48/1X,A48/)
30  FORMAT(4X,I4,5X,A48,/13X,A48)
40  FORMAT(/1X,'NOMBRE DES DONNEES:',F8.0,/,1X,'CODE DE FIN DES'
*  ', ' DONNEES:',F8.0,/,1X,'CODE DE DONNEES MANQUANTES:',
*  F8.0/)
50  FORMAT(/1X,I4,2X,8F10.2)
    END
```

```

C*****C
C
C IMPVAR : CETTE ROUTINE IMPRIME SUR PAPIER OU A L'ECRAN LES TITRES C
C PARAMETRES ET DONNEES D'UNE OU PLUSIEURS VARIABLES C
C ENREGISTREES SUR DISQUE C
C
C VARIABLES GLOBALES: IN,IND,TIT,P,DONNEE C
C
C ENTREE: IN,IOUT C
C
C ROUTINE APPELEE: TYPEAB,REDD C
C
C*****C

```

```

SUBROUTINE IMPVAR
INCLUDE 'COM1.DAT'
CHARACTER*1 CHOIX, FVAR*80
WRITE(IN,'1X','NOMBRE DES VARIABLES A IMPRIMER: ','$')
READ(IN,*)NVIM
IF (NVIM.GT.0)THEN
  K=1
  IS=1
  DO WHILE((K.LE.3).AND.(IS.EQ.1))
    IS=0
    WRITE(IN,*)'NUMEROS DES VARIABLES:'
    READ(IN,*)(IND(I),I=1,NVIM)
    DO IV=1,NVIM
      IX=IND(IV)
      IF (INIT(IX).EQ.0) THEN
        WRITE(IN,15)IX
        IS=1
      END IF
    END DO
    K=K+1
  END DO
  IF (K.GT.3)THEN
    RETURN
  ELSE

```

```

C
  WRITE(IOUT,*)'IMPRESSION VARIABLE PAR VARIABLE'
  WRITE(IN,5)
  READ(IN,'A1')CHOIX
  IF (CHOIX.EQ.'O')THEN
    WRITE(IN,*)'DONNER LE FORMAT D''IMPRESSION'
    READ(IN,'A80')FVAR
  END IF
  DO K=1,NVIM
    CALL TYPEAB(IND(K),7,IERR)
    IF (IERR.EQ.0)THEN
      WRITE(IN,1)
      WRITE(IOUT,10)IND(K),TIT,P
      NN=P(6)
      NN=NN-1
      CALL REDD(NN,XMOY,XMINI,XMAXI,ECTI,CVAR)
      WRITE(IOUT,'/,1X','DONNEES:','/')
      IF (CHOIX.EQ.'O')THEN
        WRITE(IOUT,FVAR)(DONNEE(LK),LK=1,NN)
      ELSE
        WRITE(IOUT,*)(DONNEE(LK),LK=1,NN)
      END IF
      WRITE(IOUT,20)NN,XMINI,XMAXI,XMOY,ECTI,CVAR
    END IF
  END DO

```



```
        END IF
    END DO
    CALL TYPEAB(NVA,4,IERR)
    END IF
END IF
RETURN
1  FORMAT('1')
5  FORMAT
*  (/1X,'VOUDRIEZ-VOUS INDIQUER UN FORMAT D'IMPRESSION',/1X,
*  'POUR LES DONNEES (O/N): ',%)
10 FORMAT(/1X,'VARIABLE #: ',I4,/,1X,17(' ')/1X,4(A48,/,1H ),
*  1X,'PARAMETRES:',/,1X,10F7.0,/)
15 FORMAT(/1X,'LA VARIABLE #:',I4,' N'EXISTE PAS DANS LE ',
*  'FICHIER',/)
20 FORMAT(/3X,'EFF:',I9/3X,'MIN:',F9.2/3X,'MAX:',F9.2/
*  3X,'MOY:',F9.2/3X,'ECT:',F9.2/3X,'CV: ',F9.2/)
END
```

```

C*****C
C
C   FMI: CETTE ROUTINE FORME LA MATRICE D'IMPRESSION
C
C   ENTREE: NMIN = NOMBRE MINIMUM DES DONNEES DE ENTRE TOUTES LES
C             VARIABLES
C             NV = NOMBRE DES VARIABLES
C
C             ILL = INDICE INITIAL DU VECTEUR DES DONNEES
C
C             IERR = 1 S'IL Y A ERREUR EN LA LECTURE DES VARIABLES
C             IERR = 0 S'IL N'Y A PAS D'ERREUR
C
C*****C

      SUBROUTINE FMI(NMIN,NV,IERR,NE,ILL,IL)
C   FORMATION DE LA MATRICE D'IMPRESSION
      INCLUDE 'COM1.DAT'
      DO J=1,NV
        IL=ILL
        CALL LIRTPD(IND(J),IERR)
        IF (IERR.EQ.1) THEN
          RETURN
        END IF
        I=1
        DO WHILE((I.LE.50).AND.(IL.LT.NMIN))
          IL=IL+1
          X(I,J)=DONNEE(IL)
          I=I+1
        END DO
      END DO
      RETURN
      END

```



\*\*\*\*\*  
ROUTINES POUR ACCES AUX STRUCTURES DES DONNEES  
\*\*\*\*\*

```

C*****C
C
C  TYPEAB:  CETTE ROUTINE PERMET D'EFFECTUER DES OPERATIONS SUR LES
C           FICHIERS DE TITRES,PARAMETRES ET DONNEES
C           ELLES SONT:
C
C           1  -- AJOUTER LES VARIABLES
C           2  -- CREER LES FICHIERS
C           3  -- OUVRIR LES FICHIERS
C           4  -- FERMER LES FICHIERS
C           5  -- ELIMINER LES VARIABLES
C           6  -- SUPPRIMER LES FICHIERS
C           7  -- CONSULTER LES VARIABLES
C           8  -- MODIFIER LES ITEMS D'UNE VARIABLE
C
C  ROUTINES APPELEES:  LETPD,LIRTPD
C
C  VARIABLES GLOBALES:IN,IENLT,INIT,IENLR,K5,K6,K7,K9,NOM1,NOM2,NOM3
C                     NOM4,LREGT,LREGP,LREGD,NMV
C  ENTREE:  INIT,K5,K6,K7,K9,LREGT,LREGP,LREGD,NMV
C
C           NVA = NUMERO DE VARIABLE
C           NCODE = CODE QUI PEUT PRENDRE LES VALEURS 1 A 8 POUR
C                  INDiquer LE TYPE D'OPERATION A EFFECTUER
C  SORTIE: IERR = 1 SI L'OPERATION A EU LIEU SANS ERREUR
C              = 0 SI L'OPERATION N'A PAS EU LIEU
C
C*****C

```



```

SUBROUTINE TYPEAB(NVA,NCODE,IERR)
INCLUDE 'COM1.DAT'
CHARACTER*1 CHOIX
LOGICAL TEST

```

```

GO TO(10,20,30,40,50,60,70,80,90)NCODE

```

```

C      AJOUTE D'UNE VARIABLE AUX FICHIERS DES VARIABLES
10     READ(K9'1)(IENLT(I),I=1,NMV)
      READ(K9'13)TTDIS,TEDIS,TT
      IT=TTDIS
      IF(TTDIS.EQ.0)THEN
          WRITE(IN,*)'IL N'Y A PLUS DE PLACE DANS LE FICHIER'
          RETURN
      END IF
C      ID EST LE NUMERO D'UN ENREGISTREMENT DES DONNEES DISPONIBLE
      ID=TEDIS
      READ(K9'2)(INIT(I),I=1,NMV)
      K=1
      DO I=3,12
C          LECTURE DES POINTEURS DES ENREGISTREMENTS DES DONNEES
          READ(K9'I)(IENLR(II),II=K,K+NMV-1)
          K=K+NMV
      END DO
      CALL LETPD(ISW)
      IF(ISW.EQ.0)THEN
          WRITE(IN,*)'ERREUR EN ECRITURE SUR DISQUE'
          RETURN
      END IF
      IT= TTDIS
      TTDIS=IENLT(TTDIS)

```

```

C      ENCHAINEMENT DE L'ENREGISTREMENT DE NUMERO IT A LA LISTE
C      DES TITRES ORGANISEE COMME FICHIER
      IENLT(IT)=TT
      WRITE(K9'2)(INIT(I),I=1,NMV)
      TT=IT
      K=1
      DO I=3,12
C      ECRITURE DE 100 POINTEURS DES ENREGISTREMENTS CHAQUE FOIS
      WRITE(K9'I)(IENLR(II),II=K,K+NMV-1)
      K=K+NMV
      END DO
      WRITE(K9'1)(IENLT(I),I=1,100)
      WRITE(K9'13)TTDIS,TEDIS,TT
      WRITE(IN,11)IT
11     FORMAT(1X,'LA VARIABLE EST STOCKEE DANS LA POSITION ',I4,/)
      RETURN

C      'CREATION DES FICHIERS'
20     WRITE(5,'1X','NOM DU FICHIER DES VARIABLES: ','$')
      READ(5,23)NOM0
      NOM1=NOM0//'.TIT.'
      NOM2=NOM0//'.PAR.'
      NOM3=NOM0//'.DON.'
      NOM4=NOM0//'.AUX.'
      OPEN(UNIT=K5,FILE=NOM1,STATUS='NEW',ERR=30,ACCESS='DIRECT',
*      ASSOCIATE VARIABLE=IT,RECORDSIZE=LREGT,FORM='FORMATTED',
*      CARRIAGE CONTROL='LIST')
      WRITE(IN,*)'CREATION D'UN NOUVEAU FICHIER'
      OPEN(UNIT=K6,FILE=NOM2,STATUS='NEW',ERR=30,ACCESS='DIRECT',
*      ASSOCIATE VARIABLE=IP,RECORDSIZE=LREGP)
      OPEN(UNIT=K7,FILE=NOM3,STATUS='NEW',ERR=30,ACCESS='DIRECT',
*      ASSOCIATE VARIABLE=IT,RECORDSIZE=LREGD)
      OPEN(UNIT=K9,FILE=NOM4,STATUS='NEW',ERR=30,ACCESS='DIRECT',
*      RECORDSIZE=NMV)
      CALL INPOIN
      RETURN
23     FORMAT(A5)
      RETURN

C      'OUVERTURE DES FICHIERS'
30     NE=0
      OPEN(UNIT=K5,FILE=NOM1,STATUS='OLD',ERR=31,ACCESS='DIRECT',
*      ASSOCIATE VARIABLE=IT,RECORDSIZE=LREGT,FORM='FORMATTED',
*      CARRIAGECONTROL='LIST')
      OPEN(UNIT=K6,FILE=NOM2,STATUS='OLD',ERR=31,ACCESS='DIRECT',
*      ASSOCIATE VARIABLE=IP,RECORDSIZE=LREGP)
      OPEN(UNIT=K7,FILE=NOM3,STATUS='OLD',ERR=31,ACCESS='DIRECT',
*      ASSOCIATE VARIABLE=IT,RECORDSIZE=LREGD)
      OPEN(UNIT=K9,FILE=NOM4,STATUS='OLD',ERR=31,ACCESS='DIRECT',
*      RECORDSIZE=NMV)
      RETURN
31     WRITE(IN,*)'ERREUR EN OUVERTURE DES FICHIERS'
      RETURN
33     FORMAT(A5)

```



```

C      'FERMETURE DES FICHIERS'
40      CLOSE(K5)
          CLOSE(K6)
          CLOSE(K7)
          CLOSE(K8)
          CLOSE(K9)
          RETURN

C      'ELIMINATION DES VARIABLES'
50      IF (INIT(NVA).EQ.0) THEN
          WRITE(IN,51)NVA
      ELSE
C      AJOUTER L'ENREGISTREMENT LIBERE A LA PILE DES ENREGISTREMENTS
C      DISPONIBLES DES TITRES
          IENLT(NVA)=TTDIS
          TTDIS=NVA
C      LIBERER L'ENREGISTREMENT DES DONNEES IRL A LA PILE DES
C      ENREGISTREMENTS DISPONIBLES
          IRL=INIT(NVA)
          INIT(NVA)=0
          DO WHILE (IRL.NE.0)
              IX=IENLR(IRL)
              IENLR(IRL)=TEDIS
              TEDIS=IRL
              IRL=IX
          END DO
      END IF
      WRITE(K9'1')(IENLT(I),I=1,NMV)
      WRITE(K9'2')(INIT(I),I=1,NMV)
      K=1
      DO I=3,12
C      ECRITURE DE POINTEURS DES ENREGISTREMENTS SUR DISQUE
          WRITE(K9'I')(IENLR(II),II=K,K+NMV-1)
          K=K+NMV
      END DO
      WRITE(K9'13)TTDIS,TEDIS,TT
51      FORMAT(5X,'LA VARIABLE #:',I4,' N''EXISTE PAS')
      RETURN

C      'SUPPRESSION DES FICHIERS'
60      WRITE(IN,*)'C'EST BIEN TOUT LE FICHIER QUE VOUS VOULEZ '
          WRITE(IN,'1X','SUPPRIMER (O/N) ','$')
          READ(IN,'A1')CHOIX
          IF (CHOIX.EQ.'N') THEN
              RETURN
          END IF
          NOM1=NOM0//'.TIT.'
          NOM2=NOM0//'.PAR.'
          NOM3=NOM0//'.DON.'
          NOM4=NOM0//'.AUX.'
          OPEN(UNIT=K5,FILE=NOM1,STATUS='DELETE',ACCESS='DIRECT',
*          ASSOCIATE VARIABLE=IT,RECORDSIZE=LRCGT)
          OPEN(UNIT=K6,FILE=NOM2,STATUS='DELETE',ACCESS='DIRECT',
*          ASSOCIATE VARIABLE=IP,RECORDSIZE=LREGF)
          OPEN(UNIT=K7,FILE=NOM3,STATUS='DELETE',ACCESS='DIRECT',
*          ASSOCIATE VARIABLE=IT,RECORDSIZE=LREGD)
          OPEN(UNIT=K9,FILE=NOM4,STATUS='DELETE',ACCESS='DIRECT',
*          RECORDSIZE=NMV)

```

```

      IFIN=,TRUE,
      RETURN
2     FORMAT(A5)
C     CONSULTATION DE LA VARIABLE DE NUMERO NVA SUR DISQUE
70    CALL LIRTPD(NVA,IERR)
      RETURN

C     'MODIFICATION DES VALEURS D''UNE VARIABLE'
80    WRITE(IN,*)'MODIFICATION DES TITRES? (O/N)'
      READ(IN,'A1')CHOIX
      IF (CHOIX.EQ,'O')THEN
          WRITE(IN,*)'INTRODUIR LES 4 LIGNES DES TITRES'
          DO I=1,4
              READ(IN,'A48')TIT(I)
          END DO
      END IF

      WRITE(IN,*)'MODIFICATION DES PARAMETRES? (O/N)'
      READ(IN,'A1')CHOIX
      IF (CHOIX.EQ,'O')THEN
          WRITE(IN,*)'INTRODUIR LES PARAMETRES (FORMATO LIBRE)'
          READ(IN,*)(P(I),I=1,10)
      END IF

      WRITE(IN,*)'MODIFICATION DES DONNEES?(O/N)'
      READ(IN,'A1')CHOIX
      IF (CHOIX.EQ,'O')THEN
          WRITE(IN,*)'NOMBRE DES DONNEES A MODIFIER:'
          READ(IN,*)NMODIF
          IF(NMODIF.GT,0)THEN
              DO I=1,NMODIF
                  WRITE(IN,*)
*      'DONNER L''INDICE DE LA DONNEE A MODIFIER:'
                  READ(IN,*)J
                  WRITE(IN,*)'NOUVELLE VALEUR:'
                  READ(IN,*)DONNEE(J)
              END DO
          END IF
      END IF

      GO TO 10
C     LECTURE DES TITRES
90    IERR=0
      IDR=INIT(NVA)
      IF (IDR.NE,0)THEN
          READ(K5'NVA','4A48')TIT
          IERR =1
      END IF
      RETURN
      END

```



```

C*****C
C
C LETPD : AJOUTE UNE VARIABLE AU FICHIER DE TITRES,PARAMETRES ET
C         DONNEES
C
C VARIABLES GLOBALES: IN,TIT,P,DONNEES
C ROUTINE APPELEE: LIECDV
C
C ENTREE: IN,TIT,P,DONNEES
C
C SORTIE: ISW = CODE QUI INDIQUE SI L'OPERATION S'EST EFFECTUEE
C           CORRECTEMENT
C           ISW = 0 S' IL Y A UNE ERREUR
C           = 1 SI L'OPERATION S'EST EFFECTUEE CORRECTEMENT
C*****C

```

```

SUBROUTINE LETPD(ISW)
C ECRITURE DES TITRES PARAMETRES ET DONNEES
  INCLUDE 'COM1.DAT'
  ISW=1
  IP=IT
  WRITE(IN,*)
  WRITE(IN,*)'TITRE:'
  WRITE(IN,*)
  DO IZZ=1,4
    WRITE(IN,'1X,A48')TIT(IZZ)
  END DO
  WRITE(K5'IT','4A48')TIT
  IT=IT-1
  WRITE(K6'IP,ERR=20)(P(I),I=1,10)
C ECRITURE DES DONNEES
  CALL LIECDV(ISW)
  RETURN
10  WRITE(IN,*)'ERREUR EN ECRITURE DES TITRES'
    ISW=0
    RETURN
20  WRITE(IN,*)'ERREUR EN ECRITURE DES PARAMETRES'
    ISW=0
    RETURN
END

```

```

C*****C
C      LIECDV : ECRIT LES DONNEES D'UNE VARIABLE SUR DISQUE      C
C      VARIABLES GLOBALES: IN,K7,DONNEE,TEDIS,INIT,DONNEE,IENLR  C
C      ENTREE : IN,K7,DONNEE,TEDIS,INIT,DONNEE,IENLR,LREGD      C
C      SORTIE : TEDIS,INIT,IENLR                                  C
C          ISW = CODE Q' INDIQUE SI L'OPERATION S'EST EFFECTUEE  C
C                  CORRECTEMENT                                  C
C                  = 1 S' IL N Y A PAS D'ERREUR                  C
C                  = 0 S'IL Y A ERREUR                           C
C*****C
C      SUBROUTINE LIECDV(ISW)
C      ECRITURE DES DONNEES D'UNE VARIABLE SUR DISQUE
C      INCLUDE 'COM1.DAT'
C      ISW=1
C      J=TEDIS
C      CALCUL DU NOMBRE DES ENREGISTREMENTS NECESSAIRES
C      NE=(ND/LREGD)
C      IF((ND- ND/LREGD*LREGD).GT.0)THEN
C          NE=NE+1
C      END IF
C      VERIFICATION DE L'EXISTENCE D'UN NOMBRE SUFFISSANT DES ENREGISTRE
C      DISPONIBLES
C      IC=1
C      DO I=1,NE
C          IF(J.EQ.0)THEN
C              WRITE(IN,*)
C          * 'NOMBRE DES ENREGISTREMENTS DISPONIBLES INSUFFISSANTS'
C              INDICATION D' INEXISTENCE DE REGISTRES SUFFISANTS
C              ISW=0
C              RETURN
C          END IF
C          J=IENLR(J)
C      END DO
C      IF(ISW.EQ.1)THEN
C      ECRITURE DES DONNEES SUR DISQUE EN RETIRANT DE LA LISTE DES
C      ENREGISTREMENTS DISPONIBLES ET ON FAISANT L'ENCHAINEMENT DE
C      CES ENREGISTREMENTS
C          K=1
C          J=TEDIS
C          INIT(TTDIS)=J
C          TEDIS=IENLR(TEDIS)
C          IENLR(J)=0
C          WRITE(K7'J)(DONNEE(III),III=K,K+LREGD-1)
C          K=K+LREGD
C          KK=J
C      IF(NE.GT.1)THEN
C          DO I=2,NE
C              J=TEDIS
C              TEDIS=IENLR(TEDIS)
C              IENLR(J)=0
C              IENLR(KK)=J
C              KK=J
C          WRITE(K7'J)(DONNEE(III),III=K,K+LREGD-1)
C          K=K+LREGD

```



```
END DO  
END IF  
END IF  
RETURN  
30 WRITE(IN,*)'ERREUR EN LECTURE DES DONNEES'  
ISW=0  
RETURN  
END
```

```

C*****C
C
C LIRTPD: LIT LES TITRES,PARAMETRES ET DONNEES D'UNE VARIABLE
C          ENREGISTREE SUR DISQUE
C
C VARIABLES GLOBALES:  INIT,K5,K6,K7,K9,DONNEES,NMV,IENLR,TIT
C                      P,DONNEE
C
C ENTREE:  INIT,K5,K6,K7,K9,NMV
C          NUMVAR = NUMERO DE LA VARIABLE A LIRE
C
C SORTIE:  TIT,P,DONNEE
C          IERR = 0 SI L'OPERATION S'EST EFFECTUEE CORRECTEMENT
C              = 1 SI IL Y A UNE ERREUR
C
C*****C

      SUBROUTINE LIRTPD(NUMVAR,IERR)
C      LIT LES TITRES,PARAMETRES ET DONNEES D'UNE VARIABLE DU DISQUE
      INCLUDE 'COM1.DAT'
      IERR=0
      READ(K9'2)(INIT(I),I=1,NMV)
      IDROLE=INIT(NUMVAR)
      IF(IDROLE.NE.0)THEN
        IP=NUMVAR
        READ(K5'NUMVAR,'4A48')TIT
        READ(K6'IP)(P(I),I=1,10)
        K=1
        DO I=3,12
C          LECTURE DES POINTEURS DES ENREGISTREMENTS
          READ(K9'I)(IENLR(II),II=K,K+NMV-1)
          K=K+NMV
        END DO
        READ(K9'1)(IENLT(I),I=1,NMV)
        READ(K9'13)TTDIS,TEDIS,TT
        J=INIT(NUMVAR)
        K=1
        DO IE=1,1000
          DONNEE(IE)=0
        END DO
        DO WHILE (J.NE.0)
          READ(K7'J)(DONNEE(I),I=K,K+LREGD-1)
          K=K+NMV
          J=IENLR(J)
        END DO
      ELSE
        WRITE(IN,10)NUMVAR
        IERR=1
      END IF
10  FORMAT(5X,'LA VARIABLE #:',I4,' N'EXISTE PAS DANS LE FICHIER')
      RETURN
      END

```



```

C*****C
C
C INPOIN :CETTE ROUTINE INITIALISE LES POINTEURS DES TITRES ET
C         DONNEES ET LES IMPRIME SUR LE FICHIER AUXILIAIRE
C
C VARIABLES GLOBALES: IENLT,INIT,IENLR,NMV,TTDIS,TEDIS,TT,K9
C
C ENTREE : NMV
C
C*****C

```

```

      SUBROUTINE INPOIN
      INCLUDE 'COM1.DAT'
      DO I=1,NMV-1
        IENLT(I)=I+1
        INIT(I)=0
        IENLR(I)=I+1
      END DO
      IENLT(NMV)=0
      INIT(NMV)=0
      DO I=NMV,NENR-1
        IENLR(I)=I+1
      END DO
      IENLR(NENR)=0
      TTDIS=1
      TEDIS=1
      TT=0
C      ECRITURE DES POINTEURS DES ENREGISTREMENTS DES TITRES
      WRITE(K9'1)(IENLT(I),I=1,NMV)
      WRITE(K9'2)(INIT(I),I=1,NMV)
      K=1
C      DO I=3,12
        ECRITURE DES POINTEURS DES ENREGISTREMENTS DES DONNEES
        WRITE(K9'I)(IENLR(II),II=K,K+NMV-1)
        K=K+NMV
      END DO
      WRITE(K9'13)TTDIS,TEDIS,TT
      RETURN
      END

```

\*\*\*\*\*

ROUTINES ASSOCIEES A LA REGRESSION MULTIPLE

\*\*\*\*\*



```

C*****C
C
C      RM = CETTE ROUTINE CALCULE LES PARAMETRES DE REGRESSION
C      CORRESPONDANTS A UN MODELE LINEAIRE
C
C*****C

```

```

SUBROUTINE RM
  INCLUDE 'COM1.DAT'
  WRITE(IN,*)
  WRITE(IN,*)'          REGRESSION MULTIPLE'
  WRITE(IN,*)'          ====='
  WRITE(IN,*)
  CALL REGM
  CALL TYPEAB(NVA,4,IFRR)
  IFIN=,TRUE,
  RETURN
END

```

```

SUBROUTINE REGM
  INCLUDE 'COM1.DAT'
  CHARACTER*1 CHOIX
  CHARACTER*4 F1,F2
  F1='ENTR'
  F2='SORT'
  READ (K9'2)(INIT(I),I=1,NMV)
  CALL LVDV(IS,TIN,TOUT)
  IF(IS,EQ,1)THEN
    RETURN
  END IF
  CALL LCHOIX
  DO I=1,NVD
    IDE=I

```

```

C      TROUVER LE NUMERO MINIMUM DE DONNEES ENTRE LES VARIABLES
C      INDEPENDANTES ET DEPENDANTES
C      CALL TNMD(NMIN,IDE)
C      N=NMIN
C      INITIALISATION DES VARIABLES
C      CALL IV
C      CALCUL DES PARAMETRES
C      CALL CP(NMIN,IERR,IDE)
C      IF (IERR,EQ,1)THEN
C        RETURN
C      END IF
C      IMPRESSION DES PARAMETRES DES VARIABLES DEPENDANTES ET
C      INDEPENDANTES AU CHOIX DE L'UTILISATEUR
C      CALL IPVID(IDE)

```

```

NM=NMIN-1
WRITE(IN,20)NM,TIN,TOUT
C   CALCUL DE LA REGRESSION
CALL REG(F1,F2,K,NF,DLRES,SCERES,ECTRES,RREG,TTT,TIN,TOUT)
IF (CHOIXD.EQ.'0')THEN
    J=2
C   IMPRESSION DES STATISTIQUES DERNIER PAS
    CALL ISCP(F1,J,K,NF,DLRES,SCERES,ECTRES,RREG,TTT)
END IF
C   CALCUL DES VALEURS ESTIMEES ET DES RESIDUS DE LA REGRESSION
IF(CHOIXI.EQ.'0')THEN
    WRITE(IN,*)
    WRITE(IN,5)
    WRITE(IOUT,6)
    WRITE(IOUT,10)
END IF
CALL CVS(IERR,IDE)
IF (IERR.EQ.1)THEN
    RETURN
END IF
IF (CHOIXS.EQ.'0')THEN
C   ECRIRE LA VARIABLE ESTIME SUR DISQUE
    CALL EVSD(IDE,1)
END IF
IF (CHOIXR.EQ.'0') THEN
C   ECRIRE LES RESIDUS SUR DISQUE
    CALL EVSD(IDE,2)
END IF
END DO
IFIN=,TRUE,
RETURN
5   FORMAT('1')
6   FORMAT(19X,'IMPRESSION DES VALEURS OBSERVEES ET CALCULEES',
*   //,30X,'DE LA VARIABLE DEPENDANTE'//)
10  FORMAT(15X,'NUMERO',8X,'Y',12X,'Y EST',13X,'RESIDUS'//)
20  FORMAT(//10X,'N =',I4,10X,'T1 = ',F4.2,10X,'T2 = ',F4.2,/)
END

```



```

C*****do
C
C  CCPF: CALCULE LES COEFFICIENTS DE CORRELATION PARTIELLE ET
C        LA VARIABLE F POUR TESTER SA SIGNIFICATION.
C
C  VARIABLES GLOBALES: A,R,F,N,M
C
C  ENTREE:  A,N,M
C           NVEC = NUMERO DES VARIABLES DANS L'EQUATION DE REGRESSION
C
C  SORTIE:  R,F
C*****

```

```

C      SUBROUTINE CCPF(NVEC)
C      CALCUL DES COEFFICIENTS DE CORRELATION PARTIEL ET F POUR
C      TESTER SA SIGNIFICATION
C      INCLUDE 'COM1.DAT'
C      MM=M-1
C      DO I=1,MM
C        IF(A(I,I).GT.0)THEN
C          R(I)=A(I,M)/SQRT(A(I,I)*A(M,M))
C          F(I)=(N-NVEC-2)*R(I)**2/(1-R(I)**2)
C        END IF
C      END DO
C      RETURN
C      END

```

```

C*****C
C
C  CCRCVE : CALCULE LES COEFFICIENTS DE CORRELATION PARTIELS, C
C            COVARIANCE, VARIANCE ET ECART TYPE POUR LES VARIABLES C
C            EXPLICATIVES ET LA VARIABLE DEPENDANTE A PARTIR DE LA C
C            MATRICE DES SOMMES DE CARRES ET DES SOMMES DES PRODUITS C
C            DES ECARTS C
C
C  VARIABLES GLOBALES: A,NMIN,M,R,COV,VARIA,ECT C
C
C  ENTREE: A,NMIN,M C
C
C  SORTIE: R COV,VARIA,ECT C
C*****C

```

```

C      SUBROUTINE CCRCVE(NMIN)
C      CALCUL DU COEFFICIENT DE CORRELATION R, COVARIANCE COV,
C      VARIANCE VARIA ET ECART TYPE ECTY
C      INCLUDE 'COM1.DAT'
      K=0
      DO I=1,M-1
        DO J=I+1,M
          K=K+1
          R(K)=A(I,J)/(A(I,I)*A(J,J))*0.5
          COV(K)=A(I,J)/(NMIN-2)
        END DO
        VARIA(I)=A(I,I)/(NMIN-2)
        ECT(I)=SQRT(VARIA(I))
      END DO
      VARIA(M)=A(M,M)/(NMIN-2)
      ECT(M)=SQRT(VARIA(M))
      RETURN
      END

```



```

C*****C
C
C  CS1: CALCULE LES PARAMETRES POUR L'ANALYSE DE LA VARIANCE
C
C  VARIABLES  GLOBALES: A,D,M,N,
C
C  ENTREE: NVEC=NOMBRE DES VARIABLES DANS L'EQUATION DE REGRESSION
C
C  SORTIE: DLRES =  NOMBRE DE DEGRES DE LIBERTE RESIDUELS
C           SCERES = SOMME DES CARRES DES ECARTS RESIDUELS
C           CMRES = CARRE MOYEN RESIDUEL
C           DLREG = NOMBRE DE DEGRES DE LIBERTE DE LA REGRESSION
C           SCEREG = SOMME DES CARRES DES ECARTS DE LA REGRESSION
C           ECTRES = ECART TYPE RESIDUEL
C           RREG = COEFFICIENT DE DETERMINATION MULTIPLE
C           FF = VARIABLE F DE SNEDECOR POUR TESTER LA SIGNIFICATION
C                DE LA REGRESSION
C           CMREG = CARRE MOYEN DE LA REGRESSION
C
C*****C

```

```

      SUBROUTINE CS1(NVEC,DLRES,SCERES,CMRES,DLREG,SCEREG,
*  CMREG,FF,ECTRES,RREG)
      INCLUDE 'COM1.DAT'
C  DEGRE DE LIBERTE RESIDUELLE
      DLRES=N-NVEC-1
C  SOMME DES CARRES DES ECARTS RESIDUELS
      SCERES=A(M,M)
C  CARRE MOYENNE RESIDUELLE
      CMRES=SCERES/DLRES
C  DEGRES DE LIBERTE DE LA REGRESSION
      DLREG=NVEC
C  SOMME DES CARRES DES ECARTS DE LA REGRESSION
      SCEREG=D(M)-SCERES
C  CARRE MOYEN DE LA REGRESSION
      CMREG=SCEREG/NVEC
C  F-STATISTIQUE DE LA REGRESSION
      FF=CMREG/CMRES
C  ECART TYPE RESIDUEL
      ECTRES=SQRT(CMRES)
C  CORRELATION MULTIPLE
      RREG=SCEREG/D(M)
      RETURN
      END

```

```

C*****43
C
C CS2: CETTE ROUTINE CALCULE LES COEFFICIENTS DE LA REGRESSION B(I),
C L'ERREUR STANDART S(I) POUR B(I) ET LA VALEUR F POUR TESTER
C LA SIGNIFICATION DES COEFFICIENTS DE REGRESSION
C
C VARIABLES GLOBALES: A,B,M,F,XM
C
C ENTREE: A,S,XM
C CMRES = CARRE MOYEN RESIDUEL
C SORTIE: B,S,F
C
C*****C

SUBROUTINE CS2(CMRES)
C CALCULE LES COEFFICIENTS DE LA REGRESSION B0,B(I),L'ERREUR STANDART
C S(I) POUR B(I),LE F POUR ELIMINER X(I)
INCLUDE 'COM1.DAT'
DO I=1,M
    B(I)=0
END DO
J=0
SOMME=0
DO I=1,M
    IF(A(I,I).LT.0)THEN
        B(I)=A(I,M)
        S(I)=SQRT(-A(I,I)*CMRES)
        F(I)=(B(I)/S(I))**2
        SOMME=SOMME + B(I)*XM(I)
    END IF
END DO
B0=XM(M)-SOMME
RETURN
END

```



```

C*****C
C
C  CSPE: CETTE ROUTINE  CALCULE LA MOYENNE, LE MINIMUM, LE MAXIMUM, ET
C        LA MATRICE DE SOMMES DES CARRES ET DES PRODUITS DES ECARTS.
C        POUR N DONNEES DES VARIABLES EN UTILISANT LES RESULTATS DE
C        DES CALCULS POUR N-1 DONNEES
C
C  VARIABLES GLOBALES: D,X,XM,XMAX,A,E,M
C
C  ENTREE: X,XM,XMAX,XMIN,A,D,E
C
C  SORTIE: XM,XMAX,XMIN,A,NE
C        NE=OMBRE DES DONNEES DEJA TRAITEES, IL EST=J SI IL YA MOINS
C        DE 50 ELEMENTS
C
C*****C

```

```

SUBROUTINE CSPE(NE,I)
INCLUDE 'COM1.DAT'
NE=NE+1
DO K=1,M
  D(K)=X(I,K)-XM(K)
  XM(K)=(XM(K)*(NE-1)+X(I,K))/NE
  E(K)=X(I,K)-XM(K)
  IF (X(I,K).GT.XMAX(K)) THEN
    XMAX(K)=X(I,K)
  END IF
  IF (X(I,K).LT.XMIN(K)) THEN
    XMIN(K)=X(I,K)
  END IF
END DO
DO K=1,M
  DO J=K,M
    A(K,J)=A(K,J)+D(K)*E(J)
  END DO
END DO
RETURN
END

```

```

C*****C
C
C CSPEE: CALCULE LA MOYENNE, LE MINIMUM, LE MAXIMUM, ET LA MATRICE DES
C          SOMMES DES CARRES ET DES PRODUITS DES ECARTS POUR UN MAXIMUM
C          DE 50 DONNEES PAR VARIABLE
C
C VARIABLES GLOBALES: X,P
C
C ROUTINE APPELEE: CSPE
C
C ENTREE: X,P
C          NE = NOMBRE DES DONNEES DEJA TRAITEES
C
C*****C

```

```

SUBROUTINE CSPEE(NE)
  INCLUDE 'COM1.DAT'
  I=0
  DO WHILE (I.LT.50)
    I=I+1
    IC=1
    DO J=1,M
      IF(X(I,J).EQ.P(8))THEN
        IC=2
      ELSE
        IF (X(I,J).GT.P(8))THEN
          IC=3
        END IF
      END IF
    GO TO(10,30,20)IC
  END DO
  10 CALL CSPE(NE,I)
  20 END DO
  30 CONTINUE
  RETURN
END

```



```

C*****C
C
C CVS: CALCULE LES VALEURS DE Y ESTIMEES ET LES RESIDUS DE LA REGRESSION
C
C ENTREE: L= INDICE DE LA VARIABLE DEPENDANTE
C
C SORTIE: IERR = 1 SI IL Y A ERREUR
C          IERR = 0 SI IL N'Y A PAS D'ERREUR
C
C ROUTINES APPELES: FTM,YESTTD
C
C*****

```

```

      SUBROUTINE CVS(IERR,L)
      INCLUDE 'COM1.DAT'
C      CALCUL DES VALEURS DE Y ESTIMEES
      NE=0
      ILL=0
      IL=0
C      DO WHILE(IL.LT.N)
      FORMATION DE LA MATRICE DE TRAVAIL
      CALL FMT(N,IERR,NE,ILL,IL,L)
      IF(IERR.EQ.1)THEN
        RETURN
      END IF
C      CALCUL DES VALEURS DE Y ESTIMEES POUR UN MAXIMUM DE 50
C      ELEMENTS
      CALL YESTTD(NE)
      ILL=IL
      END DO
      RETURN
      END

```

```

C*****C
C
C   EVSD: CETTE ROUTINE ENREGISTRE LA VARIABLE ESTIMEE ET LES
C           RESIDUS DE LA REGRESSION SUR DISQUE.
C
C   ROUTINE APPELEE: TYPEAB
C
C   VARIABLES GLOBALES: DONNEE,E,IN,IDEF,TIT
C
C   ENTREE:  DONNEE,E,IN,IDEF,TIT
C           I= INDICE DU VECTEUR DES VARIABLES DEPENDANTES
C
C   SORTIE:  ICODE= CODE INDIQUANT LA VARIABLE A IMPRIMER
C           SI ICODE=1, IMPRESSION DE LA VARIABLE ESTIMEE
C                   SUR DISQUE
C           SI ICODE=0, IMPRESSION DES RESIDUS DE LA REGRESSION
C                   SUR DISQUE
C
C*****C

```

```

      SUBROUTINE EVSD(I,ICODE)

      INCLUDE 'COM1.DAT'
      GO TO(10,20)ICODE
10     WRITE(IN,*)'IMPRESSION DE LA VARIABLE ESTIMEE SUR DISQUE'
      CALL TYPEAB(IDEF(I),7,IERR)
      DO I=1,N
         DONNEE(I)=Y(I)
      END DO
      WRITE(IN,*)
      * 'DONNEZ LA QUATRIEME LIGNE DU TITRE POUR LA VARIABLE ESTIMEE'
      READ(IN,'A48')TIT(4)
      CALL TYPEAB(NVA,1,IERR)
      RETURN
C     IMPRESSION DES RESIDUS SUR DISQUE
      WRITE(IN,*)
20     WRITE(IN,*)'IMPRESSION DES RESIDUS DE LA REGRESSION'
      DO I=1,N
         DONNEE(I)=E(I)
      END DO
      WRITE(IN,*)
      WRITE(IN,*)
      * 'DONNEZ LA QUATRIEME LIGNE DU TITRE POUR LES RESIDUS'
      READ(IN,'A48')TIT(4)
      WRITE(IN,*)
      CALL TYPEAB(NVA,1,IERR)
      RETURN
      END

```



```

C*****
C
C FTM: FORME LA MATRICE DE TRAVAIL AVEC 50 DONNEES DE CHAQUE
C VARIABLE EXPLICATIVE ET DEPENDANTES A PARTIR DE LA POSITION
C ILL DES DONNEES
C
C VARIABLES GLOBALES: X,DONNEE,NVI,IND,IDEP,NMIN
C
C ENTREE: IND,IDEP
C NVI = NOMBRE DES VARIABLES INDEPENDANTES
C ILL = POSITION DANS LE VECTEUR DES DONNEES A PARTIR
C DUQUEL ON FORME LA MATRICE DE TRAVAIL
C L = INDICE DE LA VARIABLE DEPENDANTE IMPLIQUEE DANS LA
C REGRESSION
C
C SORTIE: IERR = 1 S'IL Y A ERREUR DANS LA FORMATION DE
C LA MATRICE
C IERR = 0 S'IL N'Y A PAS D'ERREUR
C
C ROUTINE APPELEE: TYPEAB
C
C*****

```

```

SUBROUTINE FMT(NMIN,IERR,NE,ILL,IL,L)
INCLUDE 'COM1.DAT'
DO J=1,NVI
  IL=ILL
  CALL TYPEAB(IND(J),7,IERR)
  IF(IERR.EQ.1)THEN
    RETURN
  END IF
  I=1
  DO WHILE((I.LE.50).AND.(IL.LT.NMIN))
    IL=IL+1
    X(I,J)=DONNEE(IL)
    I=I+1
  END DO
END DO
CALL TYPEAB(IDEP(L),7,IERR)
IF(IERR.EQ.1)THEN
  RETURN
END IF
J=M
IL=ILL
I=1
DO WHILE ((I.LE.50).AND.(IL.LT.NMIN))
  IL=IL+1
  X(I,J)=DONNEE(IL)
  I=I+1
END DO
RETURN
END

```

```

C*****C
C
C  IPVID : CETTE ROUTINE IMPRIME SUR PAPIER OU A L'ECRAN, POUR
C          L'ENSEMBLE DES VARIABLES EXPLICATIVES :
C
C          -- TOUTES LES TROISIEMES LIGNES DES TITRES RELATIVES AUX
C             DIFFERENTS VARIABLES
C          -- NUMEROS, MOYENNES, MINIMUM, MAXIMUM, SCE, VARIANCE ET ET ECARTS
C             TYPES ESTIMES DES DIFFERENTES VARIABLES .
C          -- NUMEROS, SPE, COVARIANCES ET COEFFICIENTS DE CORRELATION
C             DES DIFFERENTS COUPLES DE VARIABLES.
C
C          POUR CHAQUE VARIABLE DEPENDANTE :
C
C          -- NUMERO ET TROISIEME LIGNE DE TITRE
C             NUMERO, MOYENNE, MINIMUM, MAXIMUM, SCE, VARIANCE ET ECARTS
C             TYPES
C  ROUTINE APPELEE: TYPEAB
C
C  VARIABLES GLOBALES : IN, IND, XM, XMIN, XMAX, A, VARIA, FCT, COV, R, TIT,
C                      IDEP
C  ENTREE: IN, IND, XM, XMIN, XMAX, A, VARIA, ECT, COV, R
C          IDD = INDICE DE LA VARIABLE DEPENDANTE QU'ON TRAITE
C*****C

```

```

SUBROUTINE IPVID(IIID)
C  IMPRESSION DES PARAMETRES DES VARIABLES INDEPENDANTES
C  ET DEPENDANTES AU CHOIX DE L'UTILISATEUR
  INCLUDE 'COM1.DAT'
  INTEGER OUT
  CHARACTER*1 CHOIX
  OUT=IOUT
  WRITE(IN,2)
  READ(IN,10)CHOIX
  IF (CHOIX.EQ.'0')THEN
    WRITE(IN,*)
    WRITE(OUT, '1', 1X, 'VARIABLES EXPLICATIVES', '/')
    DO IVEX=1,NVI
      CALL TYPEAB(IND(IVEX),7,IERR)
      WRITE(OUT, '1X,14,3X,A48')IND(IVEX),TIT(3)
    END DO
  END IF

```



```

WRITE(OUT,20)
DO I=1,NVI
WRITE(OUT,30)IND(I),XM(I),XMIN(I),XMAX(I),A(I,I),VARIA(I),
* ECT(I)
END DO
K=0
WRITE(OUT,35)
DO I=1,M-1
DO J=I+1,M-1
K=K+1
WRITE(OUT,40)IND(I),IND(J),A(I,J),COV(K),R(K)
END DO
K=K+1
END DO
END IF
C IMPRESSION DES PARAMETRES DES VARIABLES DEPENDANTES
WRITE(IN,3)
READ(IN,10)CHOIX
IF(CHOIX.EQ.'0')THEN
WRITE(IN,*)
WRITE(OUT,/'1',1X,'VARIABLES DEPENDANTES',/)
DO IVEX=1,NVD
CALL TYPEAB(IDEP(IVEX),7,IERR)
WRITE(OUT,'1X,I4,3X,A48')IDEP(IVEX),TIT(3)
END DO
WRITE(OUT,20)
WRITE(OUT,30)IDEP(IID),XM(M),XMIN(M),XMAX(M),A(M,M),
* VARIA(M),ECT(M)
WRITE(OUT,35)
I=1
MM=0
DO K=1,M-1
MM=MM+(M-I)
WRITE(OUT,40)IDEP(IID),IND(K),A(K,M),COV(MM),R(MM)
I=I+1
END DO
END IF
2 FORMAT(/1X,'VOUDRIEZ VOUS L'IMPRESSION DES PARAMETRES',/1X,
* 'POUR LES VARIABLES INDEPENDANTES (O/N): ',%)
3 FORMAT(/1X,'VOUDRIEZ VOUS L'IMPRESSION DES PARAMETRES POUR',
* 1X,'LA VARIABLE DEPENDANTE(O/N): ',%)
10 FORMAT(A1)
RETURN
20 FORMAT(/1X,' NO MOYENNE MINIMUM MAXIMUM',9X,'SCE',11X,
* 'VARIANCE',6X,'ECART-TYPE',/)
30 FORMAT(1X,I3,E10.2,2E11.2,2E17.8,F9.2/)
35 FORMAT(/4X,'NO',3X,'NO',9X,'SPE',15X,'COV',12X,'R'/)
40 FORMAT(1X,2I5,2E18.8,F9.4)
END

```



```

C*****C
C
C   ISCP =   IMPRIME POUR CHAQUE PHASE DE CALCUL DE LA REGRESSION . C
C           MULTIPLE : C
C           - NUMERO DE LA VARIABLE ENTRANTE OU SORTANTE ET VALEUR C
C             OBSERVEE DE LA VARIABLE T DE STUDENT C
C           - NOMBRE DE DEGRES DE LIBERTE DE LA VARIANCE RESIDUELLE, C
C             SOMME DES CARRES DES ECARTS RESIDUELS, ECART-TYPE C
C             RESIDUEL ET COEFFICIENT DE DETERMINATION C
C           - COEFFICIENTS DE REGRESSION ET TERME INDEPENDANT C
C             DE L'EQUATION DE REGRESSION C
C           - ERREURS-STANDARTS DES COEFFICIENTS DE REGRESSION C
C
C   VARIABLES GLOBALES: IN, IND, B, S, F C
C
C   ENTREE: IN, IND, B, S, F, DLRES, SCERES, ECTRES, RREG, FVAR C
C           DLRES = DEGRES DE LIBERTE RESIDUELS C
C           SCERES = SOMME DES CARRES DES ECARTS RESIDUELS C
C           ECTRES = ECART TYPE RESIDUEL C
C           RREG = COEFFICIENT DE CORRELATION MULTIPLE C
C           NF = NUMERO DE PHASE C
C           FVAR = VARIABLE CONTENANT LE FORMAT D'IMPRESSION C
C           J = 1 S'IL S'AGIT DE L'IMPRESSION DES STATISTIQUES C
C               A CHAQUE PAS DE LA REGRESSION C
C           = 0 S'IL S'AGIT DE L'IMPRESSION DES STATISTIQUES C
C               DU DERNIER PAS DE LA REGRESSION C
C *****C
C
C   SUBROUTINE ISCP(FVAR, J, K, NF, DLRES, SCERES, ECTRES, RREG, TTT)
C   INCLUDE 'COM1.DAT'
C   IMPRESSION DES PARAMETRES STATISTIQUES A CHAQUE PAS
C   DE LA REGRESSION
C   CHARACTER*4 FVAR
C   IA=IN
C   IN=IOUT
C   GO TO(5,10)J
5   WRITE(IN,20)NF,FVAR,IND(K),TTT
C   RC=RREG
10  WRITE(IN,25)DLRES,SCERES,ECTRES,RC
C   WRITE(IN,30)
C   DO I=1,M-1
C       IF(B(I).NE.0)THEN
C           TTT=SQRT(F(I))
C           WRITE(IN,35)IND(I),B(I),S(I),TTT
C       END IF
C   END DO
C   WRITE(IN,40)B0
C   DO I=1,M-1
C       IF(B(I).EQ.0)THEN
C           TTT=SQRT(F(I))
C           WRITE(IN,45)IND(I),TTT
C       END IF
C   END DO
C   IN=IA
20  FORMAT(/1X,'PHASE',I4,6X,'VARIABLE',1X,A4,'ANTE =',I4,6X,
*      'T=',F10.2//)
25  FORMAT(/1X,'DLRES',F4.0,2X,'SCERES=',E15.8,3X,'ECTRES=',F8.2,2X,
*      'R**2=',E12.4)
30  FORMAT(/4X,'NO',7X,'B',8X,'ECTB',8X,'T',/)

```



```
35  FORMAT(1X,I5,F13.6,F10.6,F10.2)
40  FORMAT(5X,'0',F13.6/)
45  FORMAT(1X,I5,23X,F10.2)
    RETURN
    END
```

74

```

C*****C
C
C IV: INITIALISE LE VECTEUR DES MOYENNESS ET LA MATRICE DES SOMMES
C   DES CARRES ET DES PRODUITS DES ECARTS
C
C VARIABLES GLOBALES: XM,XMAX,XMIN,A,M
C
C ENTREE: M
C SORTIE: XM,XMAX,XMIN
C
C*****C

```

```

SUBROUTINE IV
INCLUDE 'COM1.DAT'
DO I=1,M
  XM(I)=0
  XMAX(I)= -10.E20
  XMIN(I)= 10.E20
  DO J=1,M
    A(I,J)=0
  END DO
END DO
END

```



```

C*****C
C
C LCHOIX: LIT LES CHOIX DE L'UTILISATEUR CORRESPONDANTS A :
C      - IMPRESSION DES PARAMETRES STATISTIQUES A CHAQUE PAS DE LA
C      REGRESSION
C      - IMPRESSION DES PARAMETRES STATISTIQUES DU DERNIER PAS DE
C      LA REGRESSION
C      - STOCKAGE DES VALEURS DE Y ESTIMEES
C      - STOCKAGE DES RESIDUS
C      - IMPRESSION DES VALEURS DE Y ESTIMEES ET DE RESIDUS
C
C VARIABLES GLOBALES: IN,CHOIP,CHOIXD,CHOIXS,CHOIXR,CHOIXI,IOUT
C
C*****C

```

```

SUBROUTINE LCHOIX
INCLUDE 'COM1.DAT'
CHARACTER*1 CHOIX
WRITE(IN,'1X','CHOISIR ENTRE LES OPTIONS SUIVANTES: (O/N) ','/')
WRITE(IN,'1X','IMPRESSION DES STATISTIQUES A CHAQUE PAS: ','$')
READ(IN,10)CHOIXP
WRITE(IN,'1X','IMPRESSION DE LA DERNIERE STATISTIQUE: ','$')
READ(IN,10)CHOIXD
WRITE(IN,'1X','STOCKAGE DE LA VARIABLE DEPENDANTE ESTIMEE: ','$')
READ(IN,10)CHOIXS
WRITE(IN,'1X','STOCKAGE DES RESIDUS SUR DISQUE: ','$')
READ(IN,10)CHOIXR
WRITE(IN,'1X','IMPRESSION DES VALEURS DE Y EST. ET RESIDUS: '
* , '$')
READ(IN,10)CHOIXI
WRITE(IN,*)'VOUDRIEZ-VOUS L''IMPRESSION DES RESULTATS'
WRITE(IN,15)
READ(IN,'A1')CHOIX
IF(CHOIX.EQ.'P')THEN
    IOUT = NCP
ELSE
    IOUT = IN
END IF
10  FORMAT(A1)
15  FORMAT(1X,'A L''ECRAN OU SUR PAPIER (E/P): ','$')
RETURN
END

```

```

C*****74
C
C LVDV: CETTE ROUTINE LIT ET VALIDE LES NUMEROS DES VARIABLES QUI
C INTERVIENNENT DANS LA REGRESION MULTIPLE
C
C VARIABLES GLOBALES: INIT,IND,IDEF,IN,NVO,NVF,NVD
C
C ENTREE: INIT,IN
C
C SORTIE: IND,IDEF,NVO,NVF,NVD
C IS= 1 SI LES NUMEROS DES VARIABLES NE SONT PAS VALIDES
C IS= 0 SI L'INFORMATION INTRODUITE PAR L'UTILISATEUR EST
C CORRECTE
C*****C
C
C SUBROUTINE LVDV(IS,TIN,TOUT)
C LECTURE ET VALIDATION DES NUMEROS DES VARIABLES
C DATA IVAL/1/,K/0/
C INTEGER CODE
C LOGICAL TEST
C INCLUDE 'COM1.DAT'
C TEST=,FALSE,
C LECTURE DES VALEURS LIMITEES DE T
C WRITE(IN,5)
C READ(IN,*)TIN
C 5 FORMAT(1X,'VALEUR LIMITE DE T POUR ENTRER UNE VARIABLE',/
C * 1X,'DANS L'EQUATION: ',/$)
C WRITE(IN,10)
C 10 FORMAT(1X,'VALEUR LIMITE DE T POUR SORTIR UNE VARIABLE',/
C 1X,'DE L'EQUATION DE REGRESSION: ',/$)
C READ(IN,*)TOUT
C DO WHILE(,NOT,TEST)
C WRITE(IN,'1X','NOMBRE DES VARIABLES OBLIGATOIRES: ',/$)
C READ(IN,*)NVO
C WRITE(IN,'1X','NOMBRE DES VARIABLES FACULTATIVES: ',/$)
C READ(IN,*)NVF
C NVI=NVO+NVF
C IF((NVO,GE,0),AND,(NVF,GE,0),AND,(NVI,GT,0)) THEN
C TEST=,TRUE,
C ELSE
C WRITE(IN,*)'Erreur dans les donnees introduites..'
C WRITE(IN,*)'Recommencer..'
C END IF
C END DO
C TEST=,FALSE,
C K = 1

```



```

C      LECTURE ET TEST DE L'EXISTANCE DE VARIABLES DANS LE FICHIER DES
C      DONNEES
DO WHILE((.NOT.TEST).AND.(K.LE.3))
  IF (NVO.GT.0)THEN
    WRITE(IN,*)'NUMEROS DES VARIABLES OBLIGATOIRES(F.LIBRE)'
    READ(IN,*)(IND(I),I=1,NVO)
  END IF
  IF(NVF.GT.0)THEN
    WRITE(IN,*)'NUMEROS DES VARIABLES FACULTATIVES:'
    READ(IN,*)(IND(I),I=NVO+1,NVI)
  END IF
  WRITE(IN,'1X','NOMBRE DES VARIABLES DEPENDANTES: ','$')
  READ(IN,*)NVD
  IF(NVD.GT.0)THEN
    WRITE(IN,*)'NUMEROS DES VARIABLES DEPENDANTES:'
    READ(IN,*)(IDEP(I),I=1,NVD)
    M=NVI+1
  END IF
  K=K+1
  IS=0
  IF (NVI.GT.0)THEN
    DO IV = 1,NVI
      IF(IND(IV)).EQ.0)THEN
        WRITE(IN,15)IND(IV)
        IS=1
      END IF
    END DO
  END IF
  IF (NVD.GT.0)THEN
    DO IV = 1,NVD
      IF (IDEP(IV)).EQ.0)THEN
        WRITE(IN,15)IDEP(IV)
        IS=1
      END IF
    END DO
  END IF
  K=K+1
  IF (IS.EQ.0)THEN
    TEST = .TRUE.
  END IF
END DO
15  FORMAT(1X,'LA VARIABLE #:',I4,'N'EXISTE PAS DANS LE FICHIER')
RETURN
END

```

```

C*****C
C
C  LNOMF:CETTE ROUTINE LIT ET VALIDE LE NOM DU FICHIER DES DONNEES  C
C      CREE PAR L'UTILISATEUR                                         C
C  VARIABLES GLOBALES: IN,K8                                          C
C                                                                      C
C  ENTREE: IN,K8                                                       C
C                                                                      C
C  SORTIE: TEST : VARIABLE INDIQUANT SI LE FICHIER A ETE CREE       C
C              = .TRUE. SI LE FICHIER N'EXISTE PAS                  C
C              = .FALSE. SI LE FICHIER EXISTE                        C
C                                                                      C
C*****C
C
C      SUBROUTINE LNOMF(TEST)
C      LECTURE DU NOM DU FICHIER DES DONNEES
C      LOGICAL TEST
C      INCLUDE 'COM1.DAT'
C      TEST=.TRUE.
C      DO WHILE(TEST.AND.(I.LE.2))
C          WRITE(IN,'/1X','DONNEZ LE NOM DU FICHIER DES DONNEES: ','$')
C          READ(IN,40)NOM
C          OPEN(UNIT=K8,FILE=NOM,STATUS='OLD',ACCESS='SEQIN',
C              *   ERR=10)
C          TEST=.FALSE.
C          GO TO 20
C      10      WRITE(IN,*)'CE FICHIER N'A PAS ETE CREE'
C              I=I+1
C      20      END DO
C      40      FORMAT(A10)
C      RETURN
C      END

```



```

C*****C
C
C   REDD: CETTE ROUTINE PERMET DE TROUVER LE MINIMUM XMINI, MAXIMUM
C           XMAXI, MOYENNE XMOY, ECART TYPE ECTI ET COEFFICIENTE DE
C           VARIATION CVAR, A PARTIR DES DONNEES D'UNE VARIABLE
C
C   VARIABLE GLOBALE:   DONNEE
C
C   ENTREE: DONNEE
C           NN = NOMBRE DES DONNEES DE LA VARIABLE
C
C   SORTIE: XMOY, XMINI, XMAXI, ECTI, CVAR
C
C*****C

```

```

SUBROUTINE REDD(NN,XMOY,XMINI,XMAXI,ECTI,CVAR)
  INCLUDE 'COM1.DAT'
  SX=0.
  XMINI=10.E06
  XMAXI=-10.E06
  SXC=0.
  DO I=1,NN
    SX=SX+DONNEE(I)
    SXC=SXC + DONNEE(I)**2
    IF(DONNEE(I).GT.XMAXI)THEN
      XMAXI=DONNEE(I)
    END IF
    IF(DONNEE(I).LT.XMINI)THEN
      XMINI=DONNEE(I)
    END IF
  END DO
  XMOY=SX/NN
  ECTI=SQRT((SXC-SX**2/NN)/(NN-1))
  CVAR=ECTI/XMOY
  RETURN
END

```

```

C*****C
C
C   REG   CETTE ROUTINE :
C   - EFFECTUE LE CHOIX DES VARIABLES QUI DOIVENT INTERVENIR DANS
C     L'EQUATION DE REGRESSION
C   - CALCULE ET IMPRIME AU CHOIX DE L'UTILISATEUR DES PARAMETRES
C     STATISTIQUES A CHAQUE PAS DE LA REGRESSION
C   - CALCULE ET IMPRIME LES COEFFICIENTS DE REGRESSION.
C
C*****C

```

```

SUBROUTINE REG(F1,F2,K,NF,DLRES,SCERES,ECTRES,RREG,TTT,TIN,TOUT)
INCLUDE 'COM1.DAT'
CHARACTER*4 F1,F2
FIN=TIN**2
FOUT=TOUT**2
NVEC=0
IP=1
IVO=0
NF=0
DO I=1,M
  D(I)=A(I,I)
END DO
N=N-1
ICFIN=0
DO WHILE (ICFIN.EQ.0)
  GO TO(20,10)IF
  TROUVER LA VARIABLE POUR SORTIR DE L'EQUATION DE REGRESSION
10  CALL TVS(K,L)
  IF (L.EQ.0)THEN
    RETURN
  END IF
  V(K)=A(K,M)**2/A(K,K)
  FS = -(N-NVEC-1)*V(K)/A(M,M)
  IF (FS.GT.FOUT) THEN
20  CALL TVE(L,K,IVO)
  IF (L.EQ.0)THEN
    ICFIN=1
  ELSE
    V(K)=A(K,M)**2/A(K,K)
    FE=(N-NVEC-2)*V(K)/(A(M,M)-V(K))
    IF ((FE.LE.FIN).AND.(K.GT.NVO)) THEN
      ICFIN=1
    ELSE
      NVEC=NVEC+1
      NF=NF+1
      IP=2
      IF ((IVO.LE.NVO).AND.(NVO.GT.0)) THEN
        IP=1
      END IF
      FLAG=-1
      TTT=SQRT(FE)
    END IF
  END IF
ELSE
  FLAG=1
  NVEC=NVEC-1
  NF=NF+1
  TTT=SQRT(FS)

```



```

END IF
IF(ICFIN.EQ.0)THEN
  CALL SWP(K,FLAG)
  CALL CS1(NVEC,DLRES,SCERES,CMRES,DLREG
*   ,SCEREG,CMREG,FF,ECTRES,RREG)
  DO II=1,30
    B(II)=0
  END DO
  CALL CS2(CMRES,B0)
  CALL CCPF(NVEC)
  XN = N - 1
  XMX = N - NVEC - 1
  RREG = 1-(XN/XMX)*(1.- RREG)
  IF(CHOIXF.EQ.'0')THEN
    IF (FLAG,FR,-1)THEN
      J=1
      CALL ISCP(F1,J,K,NF,DLRES,SCERES,ECTRES,RREG,TTT)
    ELSE
      J=1
      CALL ISCP(F2,J,K,NF,DLRES,SCERES,ECTRES,RREG,TTT)
    END IF
  END IF
END IF
END IF
END DO
RETURN
END

```

```

C*****C
C
C   SWP : CETTE ROUTINE APPLIQUE L'OPERATEUR SWEEP A LA MATRICE "A"
C         APRES L'INTRODUCTION OU ELIMINATION D'UNE VARIABLE DE LA
C         EQUATION DE REGRESSION PERMETTANT DE TROUVER LES COEFFICIENTS
C         DE REGRESSION
C
C   ENTREE: K,FLAG
C   FLAG = 1 INDIQUE QU'ON A INTRODUIT LA VARIABLE D'INDICE K DANS
C           L'EQUATION DE REGRESSION
C   FLAG =-1 INDIQUE QU'ON A ELIMINE LA VARIABLE D'INDICE K DE
C           L'EQUATION DE REGRESSION
C
C*****C

```

```

SUBROUTINE SWP(K,FLAG)
  DIMENSION U(30)
  INCLUDE 'COM1.DAT'
  C=A(K,K)
  DO I=1,K
    U(I)=A(I,K)
    A(I,K)=0.0
  END DO
  DO I=K,M
    U(I)=A(K,I)
    A(K,I)=0.0
  END DO
  U(K)=FLAG
  DO I=1,M
    DO J=I,M
      A(I,J)=A(I,J)-U(I)*U(J)/C
    END DO
  END DO
  RETURN
END

```



```

C*****C
C
C TNMD: CETTE ROUTINE TROUVE LE # MINIMUM DES DONNEES PARMI TOUTES LES
C      VARIABLES INDEPENDANTES ET LA VARIABLE DEPENDANTE D'INDICE ID
C
C VARIABLES GLOBALES: IND, IDEP, P, NVI, IID
C
C ENTREE: IND, IDEP, P, NVI, IID
C      IID = INDICE DE LA VARIABLE DEPENDANTE
C
C SORTIE: IXMIN = NUMERO MINIMUM DES DONNEES
C
C*****C

```

```

SUBROUTINE TNMD(IXMIN, IID)
  INCLUDE 'COM1.DAT'
  IXMIN = 2000
  DO I=1, NVI
    CALL LIRTPD(IND(I), IERR)
    IF (IERR.EQ.1) THEN
      RETURN
    END IF
    IF (P(6).LT.IXMIN) THEN
      IXMIN=P(6)
    END IF
  END DO
  CALL LIRTPD(IDEP(IID), IERR)
  IF (IERR.EQ.1) THEN
    RETURN
  END IF
  IF (P(6).LT.IXMIN) THEN
    IXMIN=P(6)
  END IF
  RETURN
END

```

```

C*****C
C
C   TVE: CETTE ROUTINE TROUVE L'INDICE DE LA VARIABLE DONT L'ENTREE
C   DANS L'EQUATION PRODUIRA LA PLUS GRANDE REDUCTION DE LA SOMME
C   DES CARRES RESIDUELS, SI TELLE VARIABLE EXISTE
C
C   VARIABLES GLOBALES: A,D,IND,TOL,IN
C   ENTREE: A,D,IND,IN
C   IVO: INDICE DE LA VARIABLE OBLIGATOIRE A AJOUTER DANS
C   L'EQUATION DE REGRESSION, SI TELLE VARIABLE EXISTE
C
C   SORTIE: K = INDICE DE LA VARIABLE A INTRODUIRE DANS L'EQUATION DE
C   REGRESSION
C   L = 1 SI LA VARIABLE K CHECHE EXISTE
C   L = 0 SI TELLE VARIABLE N'EXISTE PAS
C
C*****C

      SUBROUTINE TVE(L,K,IVO)
C      K EST L'INDICE DE LA VARIABLE POUR LAQUEL SON ENTREE PRODUIRA LA
C      PLUS GRANDE REDUCTION DE LA SOMME DES CARRES RESIDUELS SI TELLE
C      VARIABLE EXISTE
      INCLUDE 'COM1.DAT'
      VMAX=0
      L=0
      IF((NVO.GT.0).AND.(IVO.LT.NVO))THEN
          IVO=IVO+1
          IF(A(IVO,IVO)/D(IVO).GT.TOL)THEN
              K=IVO
              L=1
          ELSE
              WRITE(IN,10)IND(IVO)
          END IF
      ELSE
          DO I=1,M-1
              IF (A(I,I)/D(I).GE.TOL)THEN
                  VAUX=A(I,M)**2/A(I,I)
                  IF (VMAX.LT.VAUX)THEN
                      VMAX=VAUX
                      K=I
                      L=1
                  END IF
              END IF
          END DO
      END IF
      10  FORMAT(1X,'LA VARIABLE OBLIGATOIRE #:',I4,
      *' NE PEUT PAS ETRE AJOUTEE'/1X
      *,'CAR SA VALEUR DE TOLERANCE EST AU DESSOUS DE LA VALEUR LIMITE')
      RETURN
      END

```



```

C*****C
C
C   TVS: CETTE ROUTINE DETERMINE L'INDICE K DE LA VARIABLE QUI SE TROUVE C
C   DANS LA FONCTION DE REGRESSION POUR LAQUELLE SON ELIMINATION C
C   PRODUIT LA PLUS PETITE AUGMENTATION DE LA SOMME DE CARRES C
C   RESIDUELS C
C
C   VARIABLES GLOBALES: A,NVO,M C
C
C   ENTREE: A,NVO C
C
C   SORTIE: K = INDICE DE LA VARIABLE CHERCHE C
C*****C

```

```

SUBROUTINE TVS(K,L)
INCLUDE 'COM1.DAT'
VMAX=- 10. E 05
L=0
DO I=1,M
  IF (A(I,I),LT,0) THEN
    IF (I,GT,NVO)THEN
      AUX=A(I,M)**2/A(I,I)
      IF (VMAX,LT,AUX)THEN
        VMAX=AUX
        K=I
        L=1
      END IF
    END IF
  END IF
END DO
RETURN
END

```

```
C*****C
C
C  YEST1D: CALCUL DE LA VALEUR DE Y ESTIMEE POUR UNE OCCURENCE DES
C          VARIABLES EXPLICATIVES
C
C  VARIABLES GLOBALES: B,Y,X,XM,E,IOUT
C
C  ENTREE: NE = INDICE DU VECTEUR Y QUI INDIQUE LA DONNEE A ESTIMER
C          I  = INDICE DE LA FILE DE LA MATRICE DE 50 DONNEES QUI
C          CONTIENT LES DONNEES A UTILISER DANS LE CALCUL DE Y
C
C*****C
```

```
C      SUBROUTINE YEST1D(NE,I)
C      CALCUL DE Y ESTIME POUR UNE DONNEE
C      INCLUDE 'COM1.DAT'
C      NE=NE+1
C      SOMME = 0
C      DO K=1,M
C          IF(B(K).NE.0)THEN
C              SOMME=SOMME+B(K)*X(I,K)
C          END IF
C      END DO
C      Y(NE)=B0+SOMME
C      E(NE)=X(I,M)-Y(NE)
C      IF (CHOIXI.EQ.'0')THEN
C          WRITE(IOUT,40)NE,X(I,M),Y(NE),E(NE)
C      END IF
40  FORMAT(16X,I4,4X,F10.3,7X,E10.3,8X,E10.3)
C      RETURN
C      END
```



```

C*****C
C
C  YESTTD: CETTE ROUTINE CALCULE LA VALEUR DE Y POUR TOUTES LES DONNEES C
C          DE LA MATRICE DE 50 ELEMENTS C
C
C  VARIABLES GLOBALES: X,P,Y C
C
C  ENTREE: NE = INDICE DE LA VARIABLE "Y" OU ON COMMENCE LES CALCULS C
C
C*****C
C
C  SUBROUTINE YESTTD(NE)
C  CALCUL DE Y ESTIME POUR TOUS LES DONNEES DE LA MATRICE DE 50 ELEMENTS
C  INCLUDE 'COM1.DAT'
C  I=0
C  DO WHILE(I.LE.50)
C    I=I+1
C    DO J=1,M
C      IC=1
C      IF(X(I,J).EQ.P(8))THEN
C        IC=2
C      ELSE
C        IF(X(I,J).GT.P(8))THEN
C          IC=3
C        END IF
C      END IF
C    END DO
C    GO TO(10,30,20)IC
C  10  CALL YEST1D(NE,I)
C  20  Y(NE)=P(9)
C  END DO
C  30  RETURN
C  END

```

```

C*****C
C
C          PROGRAMMES STATISTIQUES
C
C
C  CE PROGRAMME PERMET DE:
C
C  -- CREER DES FICHIERS PERMANENTS POUR STOCKER LES TITRES
C  PARAMETRES ET DONNEES DES VARIABLES A UTILISER
C  POUR EFFECTUER DES ANALYSES STATISTIQUES.
C  -- CHARGER DES VARIABLES A PARTIR DU TERMINAL OU DES
C  FICHIERS TEMPORELS CREES PAR L'UTILISATEUR.
C  -- CREER DE NOUVELLES VARIABLES QUI SONT LES RESULTATS DES
C  TRANSFORMATIONS DES VARIABLES DEJA EXISTENTES.
C  -- EFFECTUER LA REGRESSION MULTIPLE.
C  -- ELIMINER UNE OU PLUSIEURS VARIABLES CONTENUES DANS LES
C  FICHIERS PERMANENTS.
C  -- IMPRIMER LE RESULTAT DE LA REGRESSION ET LES RESIDUS SUR
C  DISQUE ET/OU SUR PAPIER.
C  -- IMPRIMER LA OU LES VARIABLES CONTENUES DANS LES FICHIERS
C  DES DONNEES PERMANENTS.
C
C  L'EXECUTION DE LA REGRESSION MULTIPLE REQUIERT
C  L'EXISTENCE D'UN FICHIER DES DONNEES PERMANENT CREE PAR
C  CE PROGRAMME LORS DU CHARGEMENT DES VARIABLES.
C
C*****C

```